

TARTU ÜLIKOOL
LOODUS- JA TÄPPISTEADUSTE VALDKOND
Matemaatika ja statistika instituut

Tuuli Puhkim

**Lambert'i W juhuslikud suurused ja nende
rakendamine kahjukindlustuses**

Finants- ja kindlustusmatemaatika eriala

Magistritöö (30 EAP)

Juhendaja Meelis Käärik

TARTU 2016

Lambert'i W juhuslikud suurused ja nende rakendamine kahjukindlustuses

Lühikokkuvõte. Käesoleva magistritöö eesmärk on anda ülevaade Lambert'i W funktsioonist ning uurida Lambert'i W juhuslike suuruste tekkemehhanismi, omadusi ja rakendamisvõimalusi asümmeetriliste andmete kirjeldamisel. Lambert'i W juhuslikud suurused on defineeritud iga tõenäosusjaotuse jaoks. Põgusalt käsitletakse algjaotusena eksponent- ja t -jaotust. Põhjalikumalt uuritakse Lambert'i funktsiooniga transformeeritud standardset normaaljaotust, selle omadusi võrreldakse asümmeetrilise normaaljaotusega. Töö praktilises osas sobitatakse mõlemaid vastavaid asukoha-skaala jaotusi ja Lambert'i W -ga teisendatud eksponentjaotust kahjusummadele ning analüüsitakse jaotuste sobivust. Saadud tulemusi võrreldakse levinumate kahjujaotustega. Aluseks on võetud Goerg'i (2011) poolt välja pakutud lähenemise idee, mida ta tutvustas artiklis "Lambert W random variables - a new family of generalized skewed distributions with applications to risk estimation".

Märksõnad: *asümmeetria, juhuslikud suurused, kahjukindlustus, tõenäosusjaotused.*

CERCS: P160 Statistika, operatsioonanalüüs, programmeerimine, finants- ja kindlustusmatemaatika.

Lambert W random variables and their applications in non-life insurance

Abstract. The purpose of this master's thesis is to give an overview of the Lambert W function and to discuss the properties and applications of Lambert W random variables when describing asymmetric data. Lambert W random variables are defined for all probability distributions. The situation, where input is exponential or t -distribution, is briefly introduced. Lambert W transformed standard normal distribution is studied more thoroughly and its properties are compared with skew-normal distribution. In the practical part of the thesis, the suitability of both corresponding location-scale distributions, as well as Lambert W transformed exponential distribution, to the test data is evaluated. Results are compared with common loss distributions. We refer to the proposed idea and results introduced by Goerg (2011) in "Lambert W random variables - a new family of generalized skewed distributions with applications to risk estimation".

Keywords: *asymmetry, random variables, non-life insurance, probability distributions.*

CERCS: P160 Statistics, operation research, programming, actuarial mathematics.

Sisukord

Sissejuhatus	5
1 Lambert'i W funktsioon	7
1.1 Lambert'i W funktsiooni mõiste	7
1.2 Omadused	9
1.3 Funktsiooni W diferentseerimine	10
1.4 Funktsiooni W integreerimine	11
1.5 Ritta arendamine ja lähendamine	11
2 Lambert'i W juhuslikud suurused	12
2.1 Definiitsioonid	12
2.2 Lambert'i W eksponentjaotus	15
2.3 Lambert'i W standardne normaaljaotus	17
2.3.1 Peaharu uurimine	18
2.3.2 Kõrvalharu uurimine	21
2.3.3 Lambert'i $W \times N(0, 1)$ jaotus- ja tihedusfunktsioon	22
2.3.4 Võrdlus asümmeetrilise normaaljaotusega	25
2.4 Lambert'i W t-jaotus	32
3 Jaotuste sobitamine andmetele	35
3.1 Parameetrite hindamine	35
3.2 Simulatsioonid	36
3.3 Jaotuste sobitamine kahjusummadele	38
Kokkuvõte	44
Kirjandus	45
Lisa A Töös kasutatud programmid R-s	46
A.1 Lambert'i $W \times \text{Exp}(\lambda)$ jaotusfunktsioon	46
A.2 Lambert'i $W \times \text{Exp}(\lambda)$ tihedusfunktsioon	46
A.3 Lambert'i $W \times N(0, 1)$ jaotusfunktsioon	46
A.4 Lambert'i $W \times N(0, 1)$ tihedusfunktsioon	47
A.5 Lambert'i $W \times T(0, 1, \nu)$ jaotusfunktsioon	48
A.6 Lambert'i $W \times T(0, 1, \nu)$ tihedusfunktsioon	49

Lisa B	Animeeritavad tihedused R-s	50
B.1	Lambert'i $W \times \text{Exp}(\lambda)$ tihedus	50
B.2	Lambert'i $W \times N(0, 1)$ ja asümmeetrilise normaaljaotuse tihedused	51

Sissejuhatus

Statistilises analüüsis tuleb tihti ette juhtumeid, kus andmed on peaaegu normaaljaotusega, kuid teatava asümmeetriaga. Väga sageli kerkivad esile sellised situatsioonid finants- ja kindlustusmatemaatikas, näiteks varade tulususe andmete korral. Senini on sellistes olukordades laialdaselt kasutusel asümmeetriline normaaljaotus, kus sümmeetrilise normaaljaotuse kuju deformeeritakse teatava asümmeetriaparametriga. Analoogiliselt on välja töötatud ka teisi asümmeetrilisi jaotusi (asümmeetriline t -jaotus). Lambert'i W juhuslike suurusi võib vaadelda kui üldistust, sest sisendjaotus võib olla suvaline. Asümmeetriline kuju saadakse algjaotuse deformeerimisel Lambert'i funktsiooniga.

Magistritöö eesmärgiks on tutvustada Lambert'i W funktsiooni reaalarvulisi harusid (nn pea- ja kõrvalharu) ning uurida Lambert'i W juhuslike suuruste tekkemehhanismi ja kasutamist üldiselt ning konkreetselt erinevate tõenäosusjaotuste korral. Töö põhirõhk on antud lähenemise uurimisel juhul, kui algjaotuseks on standardne normaaljaotus. Samuti kõrvutatakse seda traditsioonilise asümmeetrilise normaaljaotusega. Praktilise väljundina käsitletakse kahjukindlustuse andmeid, millele sobitatakse mõningaid antud magistritöös käsitletud asümmeetrilisi jaotusi, samuti enamlevinud tüüpilisi kahjukindlustusjaotusi. Võrreldakse, milline jaotus andmeid paremini kirjeldab ja kas Lambert'i W funktsiooniga jaotuse transformeerimine pakub paremat alternatiivi traditsioonilistele jaotustele.

Magistritöö koosneb kolmest peatükist. Esimeses osas defineeritakse Lambert'i W funktsioon, antakse lühiülevaade selle ajaloost ja põhilistest omadustest. Kuigi W funktsioon pole avaldatav elementaarfunktsioonide kaudu, näidatakse, kuidas seda on võimalik diferentseerida ja integreerida, samuti on toodud peaharu rittaarendus ning kaks meetodit funktsiooni lähendamiseks. Peatükk on kirjutatud tuginedes Brito, Fabião ja Staubyn'i (2008), Dence'i (2013) ning Corless'i jt (1996) töödele.

Teises peatükis tutvustatakse Goerg'i (2011) artiklis esitletud ideed juhuslike suuruste teisendamiseks asümmeetrilisteks Lambert'i juhuslikeks suurusteks. Nimelt, kui sisendiks on juhuslik suurus X mingist tõenäosusjaotusest F_X ja nullist erinev asümmeetriaparameter γ , siis saadakse vastavalt esitatud definitsioonidele väljundiks asümmeetriline juhuslik suurus Y Lambert'i $W \times F_X$ jaotusest. Positiivse γ korral tuletatakse Lambert'i W funktsiooni rakendades jaotus- ja tihedusfunktsioon kolme jaotusetüübi (mitteskaleeritud ja mittetsentraalse, skaleeritud ning asukoha-skaala) Lambert'i $W \times F_X$ juhusliku suuruse Y jaoks. Saadud valemeid rakendatakse eksponentjaotuse, standardse normaaljaotuse ja

tsentraalse t-jaotuse puhul. Esitatud on graafikud nende jaotus- ja tihedusfunktsioonidest erinevate asümmeetriaparaameteeride korral. Normaalkaotust on käsitletud pikemalt, analüüsidest eraldi tihedusfunktsiooni pea- ja kõrvalharu komponenti ja sellest tulenevalt tihedusfunktsiooni kuju. Lõpuks on sisse toodud ka asümmeetrilise normaalkaotuse definitsioon (Azzalini (1985)). Võrreldakse kahte lähenemist lähtuvalt momentidest ning asümmeetriakordajast.

Kolmandas peatükis käsitletakse põgusalt Lambert'i $W \times F_X$ kaotuse paraameteeride hindamist. Kaotuste kirjeldavuse võrdluseks genereeritakse teatava asümmeetriakordajaga asümmeetrilisest normaalkaotusest juhuslikke väärtusi ning sobitatakse neile Lambert'i $W \times N(\mu_X, \sigma_X)$ ja kaotust ennast ning vastupidi. Rakendusliku näitena on käsitletud kahjusummadele asukoha-skaala asümmeetrilise normaalkaotuse, Lambert'i $W \times N(\mu_X, \sigma_X)$ ja Lambert'i $W \times Exp(\lambda)$ sobitamist. Samuti sobitatakse andmetele tuntumaid kahjukindlustuskaotusi. Hinnatakse kaotuste paraameetreid ning võrreldakse kirjeldavuse headust läbi sobivusnäitajate.

Magistritöös esitatud jooniste tegemiseks ja praktilise osa läbiviimiseks on kasutatud statistikatarkvara R. Lamberti W funktsiooni arvutamiseks ning rakendusliku osa teostamiseks on kasutatud Goerg'i (2016) poolt välja töötatud lisamoodulit *LambertW*. Lisades on toodud programmid töös käsitletud kaotuste kaotus- ja tihedusfunktsioonide leidmiseks ning mõned animeeritavate tiheduste graafikute koodid. Töö on vormistatud tekstitöötlusprogrammiga L^AT_EX.

1 Lambert'i W funktsioon

Selles peatükis anname lühiülevaate Lambert'i funktsioonist, selle ajaloost ja põhilistest omadustest, tuginedes Brito jt (2008), Dence'i (2013) ning Corless'i jt (1996) töödele.

1.1 Lambert'i W funktsiooni mõiste

Lambert'i W funktsioon pärineb 18. sajandist, mil Šveitsi matemaatikud Johann Heinrich Lambert (1728-1777) ja Leonhard Paul Euler (1707-1783) arendasid trinoomvõrrandi ritta, kuid ei andnud sellele nime. Seda funktsiooni kasutati järgnevatel sajanditel korraldult erinevate probleemide lahendamisel. Funktsioon sai lõpuks nime 1980. aastate algul, kui arvutialgebra süsteemis *Maple* defineeriti see W nime all. Täpne põhjus sellise tähe valikuks pole selge. Juhtumisi andis Inglise matemaatik Edward Maitland Wright (1906-2005) olulise panuse erinevate funktsiooni W aspektide osas ning uuris esimesena funktsiooni kompleksseid väärtusi, mistõttu on tähel W lisaolulisus.

Lambert'i teosed sisaldasid palju eeltööd selle funktsiooni defineerimiseks ning kuigi ta ei defineeritud funktsiooni ennast, nimetati see tema auks. Lambert'i W funktsiooni nimetatakse mõnedes teostes ka oomega-funktsiooniks kreeka tähe Ω tiheda esinemise tõttu seonduvas kirjanduses.

Lambert'i W funktsioon on funktsiooni $f(x') = x'e^{x'}$ ($x' \in \mathbb{R}$) pöördfunktsioonide hulk. Teisiti kirjutades

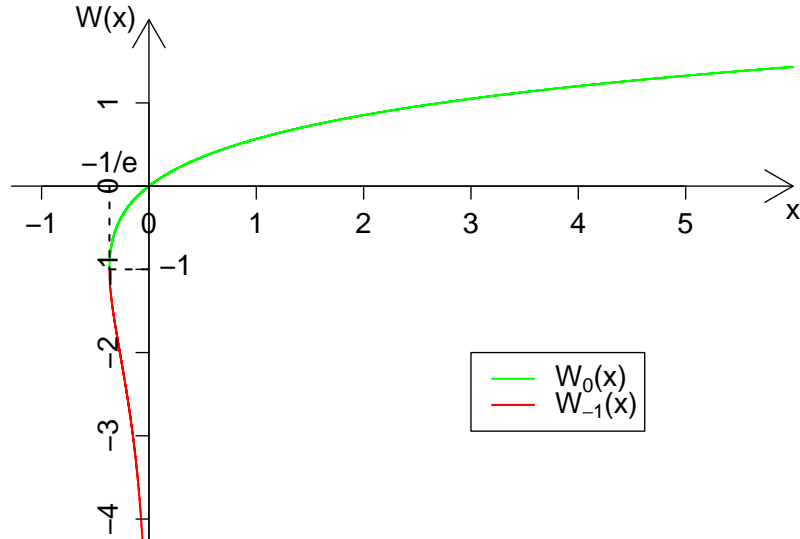
$$x' = f^{-1}(x'e^{x'}) = W(x'e^{x'}).$$

Asendades $x = x'e^{x'}$ jõuame funktsiooni W definitsioonini.

Definitsioon 1. Lambert'i W funktsioon on defineeritud kui funktsioon, mis rahuldab võrrandit

$$W(x)e^{W(x)} = x, \quad x \in \left[-\frac{1}{e}, \infty\right). \quad (1)$$

Funktsioon f pole injektiivne, seetõttu pole W üheselt määratud. Kui $x \in (-\frac{1}{e}, 0)$, siis on funktsioonil $W(x)$ kaks reaalselt väärtust. Funktsioon on üheselt määratud vahemikus $x \in [0, \infty)$. Järgime Corless'i jt (1996) kasutatud kirjaviisi. Kuna funktsioonil on mitu väärtust, siis vaatleme erinevaid harusid eraldi. Tähistame edaspidi peaharu $W_0(x)$, kus $W(x) \geq -1$, ja kõrvalharu $W_{-1}(x)$, kus $W(x) \leq -1$. Kui $x \rightarrow 0-$, siis kõrvalharu $W_{-1}(x) \rightarrow -\infty$. Peaharu kasvab aeglaselt, kuid tõkestamatult, kui $x \rightarrow \infty$ (vt joonis 1).



Joonis 1: Lambert'i W funktsioon

Üldjuhul on Lambert'i W funktsioon defineeritud nii reaalsete kui kompleksarvude jaoks ning võrrandil (1) on lõpmatu arv lahendeid, millest enamik on kompleksed. Üldjuhul tähistatakse neid $W_k(z)$, kus haru indeks $k \in \{0, \pm 1, \pm 2, \dots\}$ ja $z \in \mathbb{C}$. Selgitusi haru indekseerimise tähistuse valiku kohta saab lugeda Corless'i jt (1996) artiklist. Kui $x \in (-\infty, -\frac{1}{e})$, siis leiduvad vaid kompleksarvulised lahendid. Kui x on reaalarv, siis teised harud peale $W_0(x)$ ja $W_{-1}(x)$ on kompleksed. Kuna aga Lambert'i W juhuslikel suurustel on vaid reaalarvulised väljundid, siis piirdume selles magistritöös reaalarvulise juhuga. Detailse arutelu kompleksarvuliste harude kohta leiab Corless'i jt (1996) artiklis.

Lambert'i W funktsiooni pole võimalik avaldada elementaarfunktsioonide kaudu. Funktsioon võimaldab lahendada erinevaid eksponentfunktsiooni sisaldavaid võrrandeid. Strateegia seisneb kõigi tundmatut sisaldavate komponentide ühele võrrandi poolele viimises, et saada võrrand kujul $Y = Xe^X$, mille lahend avaldub Lambert'i funktsiooni kaudu $X = W(Y)$. Arvukates teadusharudes on Lambert'i W funktsiooni kasutuselevõtt võimaldanud tuletada suletud kujul lahendeid mudelitele, kus ilmutatud lahend pole teada. Lambert lahendas 1758. aastal trinoomvõrrandi $x = q + x^m$, arendades x ritta q astmete abil. Hiljem laiendas ta rittaarendust ka x astmete avaldamiseks. Euler teisendas Lambert'i võrrandi sümmeetrilisemale kujule

$$x^\alpha - x^\beta = (\alpha - \beta)vx^{\alpha+\beta},$$

asendades $x = x^{-\beta}$, $m = \alpha\beta$ ja $q = (\alpha - \beta)v$. Ta vaatles eriliste α ja β väärtuste juhtumeid ning jõudis uue ritta arenduseni

$$T(x) = \sum_{n \geq 1} n^{n-1} \frac{x^n}{n!}.$$

Einstein täheldas, et see rahuldab funktsionaalvõrrandit

$$T(x) = xe^{T(x)}.$$

Seega on funktsioon $T(x)$ avaldatav Lambert'i funktsiooni kaudu: $T(x) = -W(-x)$.

1.2 Omadused

Positiivsete x väärtuste korral on Lambert'i funktsiooni väärtus positiivne. Kui $x \in (-\frac{1}{e}, 0)$ siis on funktsioonil kaks negatiivset väärtust, $W_0(x)$ ja $W_{-1}(x)$. Toome mõned näited Lambert'i funktsiooni erilistest väärtustest:

- 1) $W(0) = 0$,
- 2) $W_0(-\frac{1}{e}) = W_{-1}(-\frac{1}{e}) = -1$,
- 3) $W(e) = 1$,
- 4) $W(1) = e^{-W(1)} = \ln\left(\frac{1}{W(1)}\right) = -\ln W(1) = \Omega \approx 0.5671433$.

Kuna Lambert'i funktsiooni teatakse ka oomega-funktsiooni nime all, siis sellest tulenevalt nimetatakse viimast väärtust oomega konstandiks.

Lambert'i funktsiooniga on võimalik teha arvutusi nagu teiste elementaarfunktsioonidega, kuid selleks tuleb kasutada ilmutamata funktsioonide teooriat, sest Lambert'i funktsioon pole defineeritud otsese valemiga. Uurides W graafikut, on näha, et selle asümptoodid on sarnased naturaallogaritmiga omaale, täpsemalt leiame L'Hospitali reeglit kasutades piirväärtused (vt ptk 1.3):

$$\lim_{x \rightarrow \infty} \frac{W_0(x)}{\ln x} = \lim_{x \rightarrow \infty} \frac{xW_0(x)}{x(1 + W_0(x))} = \lim_{x \rightarrow \infty} \frac{1}{\frac{1}{W_0(x)} + 1} = 1$$

ja

$$\lim_{x \rightarrow 0^-} \frac{W_{-1}(x)}{\ln(-x)} = \lim_{x \rightarrow 0^-} \frac{xW_{-1}(x)}{x(1 + W_{-1}(x))} = \lim_{x \rightarrow 0^-} \frac{1}{\frac{1}{W_{-1}(x)} + 1} = 1.$$

Samas Lambert'i W funktsiooni ja naturaalogaritmi vahe absoluutväärtus, $|W_0(x) - \ln x|$, läheb lõpmatusse, kui $x \rightarrow \infty$ (Dence (2013)).

1.3 Funktsiooni W diferentseerimine

Funktsiooni W on võimalik diferentseerida, võttes võrrandi (1) mõlemast poolest tuletise x järgi ja teisendades:

$$W'(x)e^{W(x)} + W(x)e^{W(x)}W'(x) = 1,$$

millest

$$\begin{aligned} W'(x)e^{W(x)}(1 + W(x)) &= 1 \\ \Leftrightarrow W'(x) &= \frac{e^{-W(x)}}{1 + W(x)} = \frac{W(x)}{x(1 + W(x))}, \quad x \notin \left\{0, -\frac{1}{e}\right\}. \end{aligned}$$

Võttes edasi tuletisi funktsioonist W , saab induktsiooni kasutades näidata, et n -järku tuletis avaldub kujul

$$\frac{d^n W(x)}{dx^n} = \frac{e^{-nW(x)} p_n(W(x))}{(1 + W(x))^{2n-1}}, \quad \forall n \geq 1,$$

kus polünoomid $p_n(w)$ rahuldavad rekurrentset seost

$$p_{n+1}(w) = -(nw + 3n - 1)p_n(w) + (1 + w)p'_n(w), \quad \forall n \geq 1.$$

Polünoomi algväärtus on $p_1(w) = 1$, millest edasi arvutades saame kahe järgmise polünoomi avaldisteks

$$p_2(w) = -1(w + 3 - 1) + 0(1 + w) = -w - 2,$$

$$p_3(w) = -(2w + 6 - 1)(-w - 2) + (1 + w)(-1) = 2w^2 + 8w + 9.$$

Kehtib $p_n(0) = (-n)^{n-1}$, kui $n \geq 1$.

1.4 Funktsiooni W integreerimine

Funktsiooni $W(x)$ ja seda sisaldavaid avaldisi saab integreerida, kasutades asendust $w = W(x)$, siis $x = we^w$ ja $\frac{dx}{dw} = e^w + we^w = (w+1)e^w$. Integreerime ositi kaks korda:

$$\begin{aligned}\int W(x)dx &= \int wd(we^w) = \int w(w+1)e^w dw = w(w+1)e^w - \int (2w+1)e^w dw \\ &= w(w+1)e^w - (2w+1)e^w + 2 \int e^w dw \\ &= w^2 e^w + we^w - 2we^w - e^w + 2e^w + C = (w^2 - w + 1)e^w + C \\ &= (W^2(x) - W(x) + 1)e^{W(x)} + C = x \left(W(x) - 1 + \frac{1}{W(x)} \right) + C.\end{aligned}$$

Viimasena toodud kuju on kirjanduses laiemalt levinud, kuid ei kehti juhul $x = 0$.

Analoogiliselt saab leida integraali

$$\int xW(x)dx = \int w^2(1+w)e^{2w}dw = \frac{1}{8}(2W(x)-1)(2W^2(x)+1)e^{2W(x)} + C.$$

See kehtib kõigi W harude korral, sest selle tuletise avaldisest näeme, et $\frac{d}{dw}we^w \neq 0$ ühegi haru sisepunkti korral.

1.5 Ritta arendamine ja lähendamine

Funktsiooni $W(x)$ peaharu saab arendada Taylori ritta punkti 0 ümbruses

$$W_0(x) = \sum_{n=1}^{\infty} \frac{(-n)^{n-1}}{n!} x^n = x - x^2 + \frac{3}{2}x^3 - \frac{8}{3}x^4 + \frac{125}{24}x^5 - \dots,$$

mis on koonduv $|x| < \frac{1}{e}$ korral.

Funktsiooni W saab lähendada Newtoni meetodiga (tähistame $w = W(x)$, siis $x = we^w$)

$$w_{j+1} = w_j - \frac{F(w_j)}{F'(w_j)} = w_j - \frac{w_j e^{w_j} - x}{e^{w_j} + w_j e^{w_j}} = w_j + \frac{x e^{-w_j} - w_j}{1 + w_j},$$

kus $F(w_j) = w_j e^{w_j} - x$, kuid see koondub suhteliselt aeglaselt. Kiirema ja täpsema hinnangu saavutab Halley meetodiga. Corless jt (1996) esitasid vastava iteratsiooniskeemi

$$w_{j+1} = w_j - \frac{w_j e^{w_j} - x}{e^{w_j}(1 + w_j) - \frac{(w_j+2)(w_j e^{w_j} - x)}{2w_j+2}}.$$

2 Lambert'i W juhuslikud suurused

Selles peatükis tutvustame Lambert'i W juhuslikke suurusi ja tuletame nende jaotus- ja tihedusfunktsioonid üldkujul. Hiljem uurime kolme erineva sisendjaotuse juhtusid lähemalt.

2.1 Definiitsioonid

Selles alapunktis toome Goerg'i (2011) põhjal Lambert'i W juhuslike suuruste definiitsioonid erinevate jaotusetüüpide jaoks.

Definiitsioon 2. Kui U on pidev juhuslik suurus jaotusfunktsiooniga $F_U(u) = \mathbb{P}(U \leq u)$, $u \in \mathbb{R}$, tihedusfunktsiooniga $f_U(u)$, siis ütleme, et

$$Y := U \exp(\gamma U), \quad \gamma \in \mathbb{R}, \quad (2)$$

on mittetsentraalne, mitteskaleeritud Lambert'i $W \times F_U$ juhuslik suurus asümmeetriaparametriga γ .

Parameeter γ võib omandada kõiki väärtusi reaalteljel. Kuna eksponentfunktsiooni väärtus on alati positiivne, siis on juhuslikud suurused U ja Y sama märgiga. Kui $\gamma = 0$, siis $\exp(\gamma U) = 1$ ning teisendus (2) taandub samasuseks $U \equiv Y$. Kui $\gamma > 0$, siis positiivsete U väärtuste korral faktor $\exp(\gamma U) > 1$ võimendab U väärtusi ning negatiivsete korral $0 < \exp(\gamma U) < 1$ vähendab juhusliku suuruse U mõju. Seega on positiivse asümmeetriaparametri korral Y paremale kaldu. Analoogiliselt on Y juhul $\gamma < 0$ vasakule kaldu.

Kui sisendiks on skaleeritud juhuslik suurus X , peab see omadus üle kanduma ka Lambert'i $W \times F_X$ juhuslikule suurusele. Sama kehtib asukoha-skaala jaotusfunktsiooniga juhuslike suuruste korral. Näiteks tohib sisendi erinev mõõtühik skaleerida vaid väljundit vastavalt, muud omadused, nagu asümmeetriakordaja väärtus, peavad jääma samaks. Teisendusel (2) seda omadust aga pole. Selleks, et sisendjaotusest konstrueerida asümmeetriline jaotus, mis on mastaabi- ja nihkeinvariantne (näiteks normaaljaotus), on vaja esitada veel kaks definiitsiooni.

Definiitsioon 3. Kui X on pidev juhuslik suurus skaleeritud jaotusfunktsiooniga $F_X(x|\beta)$, kus β on F_X parameeter(vektor) ja σ_X on standardhälve, siis $U = \frac{X}{\sigma_X}$, $DU = 1$, ning

ütlemel, et

$$Y := \{U \exp(\gamma U)\} \sigma_X = X \exp\left(\gamma \frac{X}{\sigma_X}\right), \quad \gamma \in \mathbb{R}, \quad \sigma_X > 0, \quad (3)$$

on skaleeritud Lambert'i $W \times F_X$ juhuslik suurus parameetervektoriga $\theta = (\beta, \gamma)$.

Definitsioon 4. Kui X on juhuslik suurus asukoha-skaala jaotusfunktsiooniga $F_X(x|\beta)$, kus β on F_X parameeter(vektor), μ_X on keskväärutus ja σ_X standardhälve, siis $U = \frac{X - \mu_X}{\sigma_X}$, $EU = 0$, $DU = 1$, ning ütlemel, et

$$Y := \{U \exp(\gamma U)\} \sigma_X + \mu_X, \quad \gamma \in \mathbb{R}, \quad \sigma_X > 0, \quad (4)$$

on asukoha-skaala Lambert'i $W \times F_X$ juhuslik suurus parameetervektoriga $\theta = (\beta, \gamma)$.

Konkreetsel sisendjaotuse puhul tuleb valida definitsioonides toodud teisenduste (2), (3), (4) vahel vastavalt sellele, mis jaotuste perekonda see kuulub. Antud magistritöös eeldame edaspidi, et asümmeetriaparameter γ on positiivne. Negatiivse korral on tuletuskäigud analoogilised.

Olgu X skaleeritud mittenegatiivne juhuslik suurus jaotusfunktsiooniga F_X ja tihedusfunktsiooniga f_X ning olgu selle standardhälve σ_X . Lambert'i funktsioon on selles piirkonnas üheselt määratud ning monotoonne. Kasutame definitsioonis 3 defineeritud juhuslikku suurust Y . Tähistame $z = \frac{y}{\sigma_X}$. Lambert'i $W \times F_X$ jaotusfunktsiooni F_Y saab leida F_X kaudu järgmiselt

$$\begin{aligned} F_Y(y) &= \mathbb{P}(Y \leq y) = \mathbb{P}(U \exp(\gamma U) \sigma_X \leq y) = \mathbb{P}(\gamma U \exp \gamma U \leq \gamma z) \\ &= \mathbb{P}(\gamma U \leq W(\gamma z)) = \mathbb{P}\left(U \leq \frac{W(\gamma z)}{\gamma}\right) = F_X\left(\frac{W(\gamma z)}{\gamma} \sigma_X\right). \end{aligned}$$

Saame skaleeritud Lambert'i $W \times F_X$ juhusliku suuruse Y jaotusfunktsiooniks

$$F_Y(y|\beta, \gamma) = F_X\left(\frac{W(\gamma y/\sigma_X)}{\gamma} \sigma_X \middle| \beta\right). \quad (5)$$

Võttes avaldisest (5) tuletise y järgi, saame leida juhusliku suuruse Y tihedusfunktsiooni

$$\begin{aligned} f_Y(y|\beta, \gamma) &= f_X\left(\frac{W(\gamma y/\sigma_X)}{\gamma} \sigma_X \middle| \beta\right) \left(\frac{W(\gamma y/\sigma_X)}{\gamma} \sigma_X\right)' \\ &= f_X\left(\frac{W(\gamma y/\sigma_X)}{\gamma} \sigma_X \middle| \beta\right) W'\left(\frac{\gamma y}{\sigma_X}\right) \\ &= f_X\left(\frac{W(\gamma y/\sigma_X)}{\gamma} \sigma_X \middle| \beta\right) \frac{\exp(-W(\gamma y/\sigma_X))}{1 + W(\gamma y/\sigma_X)}. \end{aligned} \quad (6)$$

Selleks, et saada mittenegatiivselt määratud mittetsentraalse, mitteskaleeritud juhusliku suuruse (definiitsioon 2) jaotus- ja tihedusfunktsiooni avaldised, tuleb valemities (5) ja (6) võtta skaalaparameter $\sigma_X = 1$.

Leiame nüüd analoogiliselt asukoha-skaala juhusliku suuruse jaotus- ja tihedusfunktsiooni. Olgu X asukoha-skaala juhuslik suurus jaotusfunktsiooniga F_X ja tihedusfunktsiooniga f_X ning olgu selle standardhälve σ_X , keskväärus μ_X . Definiitsioonis 4 defineeritud juhuslik suurus Y on Lambert'i $W \times F_X$ jaotusega, tähistame $z = \frac{y - \mu_X}{\sigma_X}$. Kuna nüüd pole X mittenegatiivne, siis meenutades, et Lambert'i funktsioon $W(\gamma z)$ on vahemikus $z \in \left(-\frac{1}{\gamma e}, 0\right)$ kaheselt määratud, tuleb vastavas vahemikus kasutada juhusliku suuruse Y jaotusfunktsiooni F_Y ja tihedusfunktsiooni f_Y arvutamisel nii pea- kui kõrvalharu väärtusi. Selle vahemiku jaoks saame jaotusfunktsiooniks

$$\begin{aligned} F_Y(y) &= \mathbb{P}(Y \leq y) = \mathbb{P}(U \exp(\gamma U) \sigma_X + \mu_X \leq y) = \mathbb{P}(\gamma U \exp \gamma U \leq \gamma z) \\ &= \mathbb{P}(W_{-1}(\gamma z) \leq \gamma U \leq W_0(\gamma z)) = \mathbb{P}\left(\frac{W_{-1}(\gamma z)}{\gamma} \leq \frac{X - \mu_X}{\sigma_X} \leq \frac{W_0(\gamma z)}{\gamma}\right) \\ &= \mathbb{P}\left(\frac{W_{-1}(\gamma z)}{\gamma} \sigma_X + \mu_X \leq X \leq \frac{W_0(\gamma z)}{\gamma} \sigma_X + \mu_X\right) \\ &= F_X\left(\frac{W_0(\gamma z)}{\gamma} \sigma_X + \mu_X\right) - F_X\left(\frac{W_{-1}(\gamma z)}{\gamma} \sigma_X + \mu_X\right). \end{aligned}$$

Tihedusfunktsiooni selles vahemikus saab leida analoogiliselt tuletuskäigule (6), võttes saadud jaotusfunktsiooni avaldise mõlema haru komponendist tuletise y järgi.

Kuna funktsioon $W(\gamma z)$ pole määratud piirkonnas, kus $z < -\frac{1}{\gamma e}$ ($y < -\frac{\sigma_X}{\gamma e} + \mu_X$), siis selles vahemikus $F_Y(z) = f_Y(z) = 0$. Kuna $W'(\gamma z)$ pole aga määratud punktis $z = -\frac{1}{\gamma e}$, siis on ka seal $f_Y(z) = 0$ ning tihedus pole pidev selles punktis. Piirkonnas $z \geq 0$ ($y \geq \mu_X$) on $W(\gamma z)$ üheselt määratud ja monotoonne, seega tuleb kasutada vaid peaharu väärtusi. Saime asukoha-skaala Lambert'i $W \times F_X$ juhusliku suuruse jaotusfunktsiooniks

$$F_Y(y|\beta, \gamma) = \begin{cases} 0, & \text{kui } y \leq -\frac{\sigma_X}{\gamma e} + \mu_X; \\ F_X\left(\frac{W_0(\gamma z)}{\gamma} \sigma_X + \mu_X \mid \beta\right) - F_X\left(\frac{W_{-1}(\gamma z)}{\gamma} \sigma_X + \mu_X \mid \beta\right), & \text{kui } -\frac{\sigma_X}{\gamma e} + \mu_X < y < \mu_X; \\ F_X\left(\frac{W_0(\gamma z)}{\gamma} \sigma_X + \mu_X \mid \beta\right), & \text{kui } y \geq \mu_X, \end{cases} \quad (7)$$

kus $z = \frac{y - \mu_X}{\sigma_X}$. Punktis $z = -\frac{1}{\gamma e}$ on funktsioon pidev, sest $W_0(\gamma z) = W_{-1}(\gamma z)$. Punktis $z = 0$ ($y = \mu_X$) koondub kõrvalharu komponent nulli, seega on jaotusfunktsioon pidev ka

selles punktis. Tihedusfunktsioon on kujul

$$f_Y(y|\beta, \gamma) = \begin{cases} 0, & \text{kui } y \leq -\frac{\sigma_X}{\gamma e} + \mu_X; \\ f_X\left(\frac{W_0(\gamma z)}{\gamma}\sigma_X + \mu_X \middle| \beta\right) W'_0(\gamma z) \\ \quad - f_X\left(\frac{W_{-1}(\gamma z)}{\gamma}\sigma_X + \mu_X \middle| \beta\right) W'_{-1}(\gamma z), & \text{kui } -\frac{\sigma_X}{\gamma e} + \mu_X < y < \mu_X; \\ f_X\left(\frac{W_0(\gamma z)}{\gamma}\sigma_X + \mu_X \middle| \beta\right) W'_0(\gamma z), & \text{kui } y \geq \mu_X, \end{cases} \quad (8)$$

kus $z = \frac{y - \mu_X}{\sigma_X}$ ja

$$W'(\gamma z) = \frac{\exp(-W(\gamma z))}{1 + W(\gamma z)} = \frac{W(\gamma z)}{\gamma z(1 + W(\gamma z))}. \quad (9)$$

Jaotus- ja tihedusfunktsiooni $\gamma < 0$ korral saab leida analoogiliselt. Kui $\gamma = 0$, siis on selge, et $F_Y(y|\beta, \gamma) = F_X(y|\beta)$.

Vaatleme nüüd Lambert'i W juhuslike suuruste olulisemaid tõenäosusjaotusi lähemalt.

2.2 Lambert'i W eksponentjaotus

Olgu X eksponentjaotusega juhuslik suurus parameetriga $\beta = \lambda > 0$. Tähistame seda $X \sim \text{Exp}(\lambda)$. Meenutame, et siis on juhusliku suuruse X jaotusfunktsioon kujul

$$F_X(x|\lambda) = 1 - \exp(-\lambda x), \quad \lambda > 0, x \geq 0,$$

ja tihedusfunktsioon kujul

$$f_X(x|\lambda) = \lambda \exp(-\lambda x), \quad \lambda > 0, x \geq 0.$$

Lambert'i W funktsiooniga transformeerimisel eeldame taas, et asümmeetriaparameter $\gamma > 0$. Selliselt valitud γ pakub just kindlustusvaldkonnas huvi, kuna muudab raskemaks eksponentjaotuse parema saba, mis on muidu kahjude suuruste kirjeldamiseks liiga kerge. Negatiivne asümmeetriaparameter muudaks saba veelgi kergemaks. Kuna sisendjaotus on skaalajaotuste perekonnast ning mittenegatiivselt määratud, siis on Lambert'i $W \times \text{Exp}(\lambda)$ jaotusest juhusliku suuruse Y jaotus- ja tihedusfunktsioon leitav valemite (5) ja (6) abil. Teame, et eksponentjaotusega juhusliku suuruse standardhälve on $\sigma_X = \frac{1}{\lambda}$. Saame nüüd

juhusliku suuruse Y jaotusfunktsiooniks

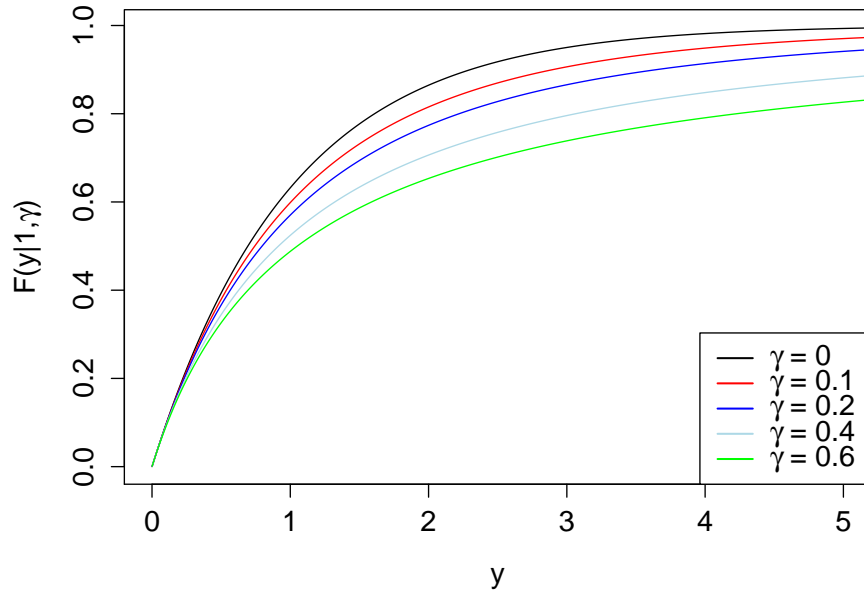
$$F_Y(y|\lambda, \gamma) = 1 - \exp\left(-\frac{W_0(\gamma\lambda y)}{\gamma}\right), \quad \lambda > 0, y \geq 0,$$

ja tihedusfunktsiooniks

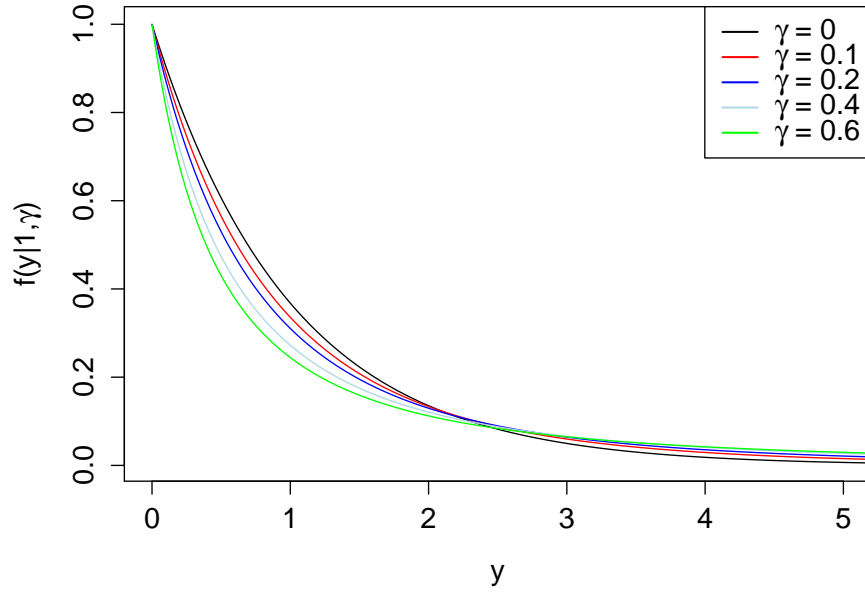
$$f_Y(y|\lambda, \gamma) = \lambda \exp\left(-\frac{W_0(\gamma\lambda y)}{\gamma}\right) \frac{\exp(-W_0(\gamma\lambda y))}{1 + W_0(\gamma\lambda y)}, \quad \lambda > 0, y \geq 0.$$

Vaatame neid funktsioone erinevate asümmeetriaparametrite korral ka joonisel. Valime eksponentjaotuse parameetriks $\lambda = 1$, siis standahälve on samuti $\sigma_X = 1$, ning asümmeetriaparametrite väärtusteks $\gamma \in \{0, 0.1, 0.2, 0.4, 0.6\}$.

Näeme, et eksponentjaotuse tihedusfunktsiooni saba muutub asümmeetriaparametri suurenedes raskemaks (vt joonis 3). Parema ettekujutuse Lambert'i $W \times \text{Exp}(\lambda)$ muutumisest skaala- ja asümmeetriaparametri muutudes saab siis, kui kasutada Rööp (2015) põhjal tehtud koodi tiheduse animeerimiseks, mis on toodud lisas B.1.



Joonis 2: Lambert'i $W \times \text{Exp}(1)$ jaotusfunktsioon erinevate asümmeetriaparametri väärtuste korral



Joonis 3: Lambert'i $W \times \text{Exp}(1)$ tihedusfunktsioon erinevate asümmeetriaparametri väärtuste korral

2.3 Lambert'i W standardne normaaljaotus

Normaaljaotus on statistikas üks laialdasemalt kasutatav jaotus, kuid tihti pole reaalsed andmed täpselt sümmeetrilised. Sellisel juhul pakub Lambert'i W võimaluse selliste andmete käsitlemiseks. Olgu X standardse normaaljaotusega juhuslik suurus, st $X \sim N(0, 1)$ ja juhusliku suuruse X parameetervektor β on $(\mu_X, \sigma_X) = (0, 1)$. Jaotusfunktsioon avaldub kujul

$$F_X(x|\beta) = \Phi(x), \quad x \in \mathbb{R}.$$

Juhusliku suuruse X tihedusfunktsioon

$$f_X(x|\beta) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right), \quad x \in \mathbb{R}.$$

Toome jaotusele juurde asümmeetriaparametri γ ja võtame appi Lambert'i W juhuslikud suurused, mille sisendjaotuseks on standardne normaaljaotus. Uurime esmalt, kuidas käitub Lambert'i $W \times N(0, 1)$ tihedusfunktsiooni peaharu komponent erinevate asümmeetriaparametrite puhul. Hiljem analüüsime analoogiliselt ka kõrvalharu.

2.3.1 Peaharu uurimine

Selles alapunktis analüüsime Lambert'i $W \times N(0, 1)$ juhusliku suuruse Y tihedusfunktsiooni peaharu komponenti. Kõrvalharu väärtusi ignoreerides saame valemit (8) kasutades leida Lambert'i $W \times N(0, 1)$ juhusliku suuruse Y tihedusfunktsiooni peaharu korral

$$\frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(W_0(\gamma y))^2}{2\gamma^2}\right) \frac{\exp(-W_0(\gamma y))}{1 + W_0(\gamma y)}, \quad (10)$$

mis on määratud piirkonnas $y > -\frac{1}{\gamma e}$.

Analüüsime saadud avaldise kuju. Uurime, milliste $\gamma > 0$ väärtuste korral leiduvad funktsioonil ekstreerumid ja milliste korral mitte. Selleks võtame avaldisest (10) tuletise y järgi ning võrdsustame nulliga. Kuna järgnev analüüs kehtib mõlema haru korral kuni muutumispäirakonna tingimuseni ning järgmises alapunktis kasutame seda sama analüüsi ka kõrvalharu jaoks, siis kasutame siinkohal ja edaspidi sarnastes situatsioonides haru eristamata tähist W . Esmalt leiame

$$\begin{aligned} \left(\frac{\exp\left(-\frac{(W(\gamma y))^2}{2\gamma^2} - W(\gamma y)\right)}{1 + W(\gamma y)} \right)' &= \frac{\exp\left(-\frac{(W(\gamma y))^2}{2\gamma^2} - W(\gamma y)\right) \left(-\frac{(W(\gamma y))^2}{2\gamma^2} - W(\gamma y)\right)'}{(1 + W(\gamma y))} \\ &\quad - \frac{\exp\left(-\frac{(W(\gamma y))^2}{2\gamma^2} - W(\gamma y)\right) (1 + W(\gamma y))'}{(1 + W(\gamma y))^2}, \end{aligned} \quad (11)$$

kus kasutades valemit (9) saame

$$\begin{aligned} \left(-\frac{(W(\gamma y))^2}{2\gamma^2} - W(\gamma y) \right)' &= -\frac{2W(\gamma y)W'(\gamma y)\gamma}{2\gamma^2} - \gamma W'(\gamma y) \\ &= -\frac{W(\gamma y) \exp(-W(\gamma y))}{\gamma(1 + W(\gamma y))} - \frac{\gamma \exp(-W(\gamma y))}{1 + W(\gamma y)} \\ &= \frac{-\exp(-W(\gamma y))(W(\gamma y) + \gamma^2)}{\gamma(1 + W(\gamma y))} \end{aligned}$$

ning

$$(1 + W(\gamma y))' = \gamma W'(\gamma y) = \frac{\gamma \exp(-W(\gamma y))}{1 + W(\gamma y)},$$

millest

$$\begin{aligned} &\left(\frac{-\exp(-W(\gamma y))(W(\gamma y) + \gamma^2)}{\gamma(1 + W(\gamma y))} \right) (1 + W(\gamma y)) - \frac{\gamma \exp(-W(\gamma y))}{1 + W(\gamma y)} \\ &= \frac{-\exp(-W(\gamma y))((W(\gamma y))^2 + (1 + \gamma^2)W(\gamma y) + 2\gamma^2)}{\gamma(1 + W(\gamma y))} \end{aligned}$$

Asendame saadud tulemused võrrandisse (11) ja lahendame

$$\frac{\exp\left(-\frac{(W(\gamma y))^2}{2\gamma^2} - W(\gamma y)\right) (-\exp(-W(\gamma y)))((W(\gamma y))^2 + (1 + \gamma^2)W(\gamma y) + 2\gamma^2)}{\gamma(1 + W(\gamma y))^3} = 0. \quad (12)$$

Võrdus (12) kehtib parajasti siis, kui lugeja võrdub nulliga ja nimetaja mitte. Kuna $\gamma > 0$, siis saame nimetajast y jaoks kitsenduse $y \neq -\frac{1}{\gamma e}$. Lugejas näeme, et kuna eksponent-funktsiooni väärtus on alati positiivne, jääb meil üle lahendada ruutvõrrand

$$(W(\gamma y))^2 + (1 + \gamma^2)W(\gamma y) + 2\gamma^2 = 0 \quad (13)$$

$W(\gamma y)$ suhtes. Selle võrrandi lahendid on kujul

$$W(\gamma y) = -\frac{1 + \gamma^2}{2} \pm \sqrt{\left(\frac{1 + \gamma^2}{2}\right)^2 - 2\gamma^2} = \frac{-1 - \gamma^2 \pm \sqrt{\gamma^4 - 6\gamma^2 + 1}}{2}. \quad (14)$$

Uurime lähemalt ruutjuure all olevat avaldist. Võrrand (14) lahendub üheselt juhul kui $\gamma^4 - 6\gamma^2 + 1 = 0$. Kahe reaalarvulise lahendi leidumiseks peab aga kehtima $\gamma^4 - 6\gamma^2 + 1 > 0$. Leiame esmalt nullkohad

$$\gamma^4 - 6\gamma^2 + 1 = 0 \Leftrightarrow \gamma^2 = 3 \pm 2\sqrt{2} \Leftrightarrow \gamma = \pm\sqrt{3 \pm 2\sqrt{2}} = \pm(\sqrt{2} \pm 1).$$

Seega nullkohti on neli. Meie uurime juhtu kus $\gamma > 0$, seega tähtsust omavad hetkel lahendid $\gamma^{(1)} = \sqrt{2} - 1 \approx 0.4142$ ja $\gamma^{(2)} = \sqrt{2} + 1 \approx 2.4142$.

Saime, et Lambert'i $W \times N(0, 1)$ juhusliku suuruse tihedusfunktsioonil on kaks ekstreemumit asümmeetriaparametri väärtuste korral piirkondadest $\gamma > \sqrt{2} + 1$ ja $0 < \gamma < \sqrt{2} - 1$. Kui $\sqrt{2} - 1 < \gamma < \sqrt{2} + 1$, siis võrrandil (14) reaalarvulised lahendid puuduvad. Jääb veel üle kontrollida, kas antud vahemikud on kooskõlas peaharu komponendi määramispiirkonnaga ehk kas saadud $W(\gamma y)$ väärtuste korral kehtib $y > -\frac{1}{\gamma e}$. Kasutades Lambert'i W funktsiooni definitsiooni (1) punktis $x = \gamma y$ tuleb lahendada võrratus

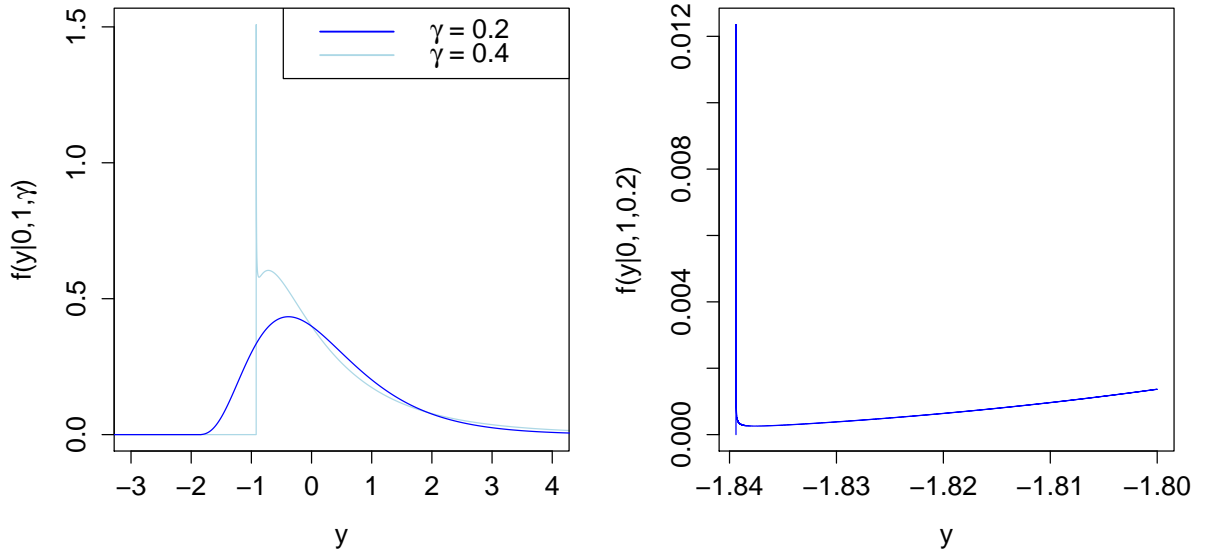
$$\frac{-1 - \gamma^2 \pm \sqrt{\gamma^4 - 6\gamma^2 + 1}}{2\gamma} \exp\left(\frac{-1 - \gamma^2 \pm \sqrt{\gamma^4 - 6\gamma^2 + 1}}{2}\right) > -\frac{1}{\gamma e}.$$

Tulemuseks saame samad piirkonnad $\gamma > \sqrt{2} + 1$ ja $0 < \gamma < \sqrt{2} - 1$. Kuna hetkel käsitleme vaid peaharu, siis kontrollime, millised γ väärtused jäävad peaharu muutumispiirkonda ehk milliste asümmeetriaparametrite korral $W(\gamma y) > -1$. Võrrandi (14) põhjal saame

võrratuse

$$\frac{-1 - \gamma^2 \pm \sqrt{\gamma^4 - 6\gamma^2 + 1}}{2} > -1 \Leftrightarrow \gamma^2 - 1 < \pm \sqrt{\gamma^4 - 6\gamma^2 + 1},$$

kus positiivsete γ väärtuste korral lahendiks on $0 < \gamma \leq \sqrt{2} - 1$. Seega on Lambert'i $W \times N(0, 1)$ juhusliku suuruse Y tihedusfunktsiooni peaharu komponendil kaks ekstreemumpunkti, kui $\gamma \in (0, \sqrt{2} - 1)$. Vaatame võrduse (12) vasakul pool toodud peaharu komponendi tuletist juhul, kui $\gamma > \sqrt{2} - 1$. Kuna siis võrrandil (13) lahendid puuduvad ning selle vasak pool on positiivne ja samuti nimetajas tegur $(1 + W(\gamma y))^3$ on positiivne, siis on tuletise avaldis negatiivne. Saime, et kui $\gamma > \sqrt{2} - 1$, siis jaotus on ühe sabaga ja kahaneb monotoonselt piirkonnas $y > -\frac{1}{\gamma e}$. Seda seaduspära on näha ka joonisel. Samuti võib täheldada, et mida väiksem on asümmeetriaparameter, seda vähem on teine ekstreemumpunkt silmaga märgatav (vt joonis 4, vasakul). Seda on näha, arvutades tihedusfunktsiooni väärtused tihedamas vahemikus ja vaadates joonist lähemalt punkti $-\frac{1}{\gamma e}$ parempoolses ümbruses (vt joonis 4 parametri $\gamma = 0.2$ jaoks).



Joonis 4: Vasakul: Lambert'i $W \times N(0, 1)$ tihedusfunktsiooni peaharu komponent erinevate asümmeetriaparametri väärtuste korral; Paremal: Lambert'i $W \times N(0, 1)$ tihedusfunktsiooni peaharu komponent $\gamma = 0.2$ korral (lõigus $[-\frac{1}{0.2e}, -1.8]$)

Näeme, et tihedusfunktsiooni peaharu komponent hakkab kasvama, kui y läheneb punktile $-\frac{1}{\gamma e}$ paremalt. Vaatame, mis selles protsessis täpsemalt juhtub. On lihtne näha, et $W_0(\gamma y) \rightarrow -1$, $(W_0(\gamma y))^2 \rightarrow 1$ ja $1 + W_0(\gamma y)$ läheb nulli paremalt poolt. Kui nüüd vaatame

tihedusfunktsiooni osa $\frac{\exp(-W_0(\gamma y))}{1 + W_0(\gamma y)}$, siis näeme, et selle väärtused lähevad lõpmatusse. Sellest tulenevalt, kui $y \rightarrow -\frac{1}{\gamma_e} +$, siis peaharu komponent (10) koondub lõpmatusse.

2.3.2 Kõrvalharu uurimine

Uurime nüüd eelmise alapunktiga analoogiliselt tihedusfunktsiooni kõrvalharu komponenti määramispiirkonnas $\left(-\frac{1}{\gamma_e}, 0\right)$. Kasutades valemit (8) on see kujul

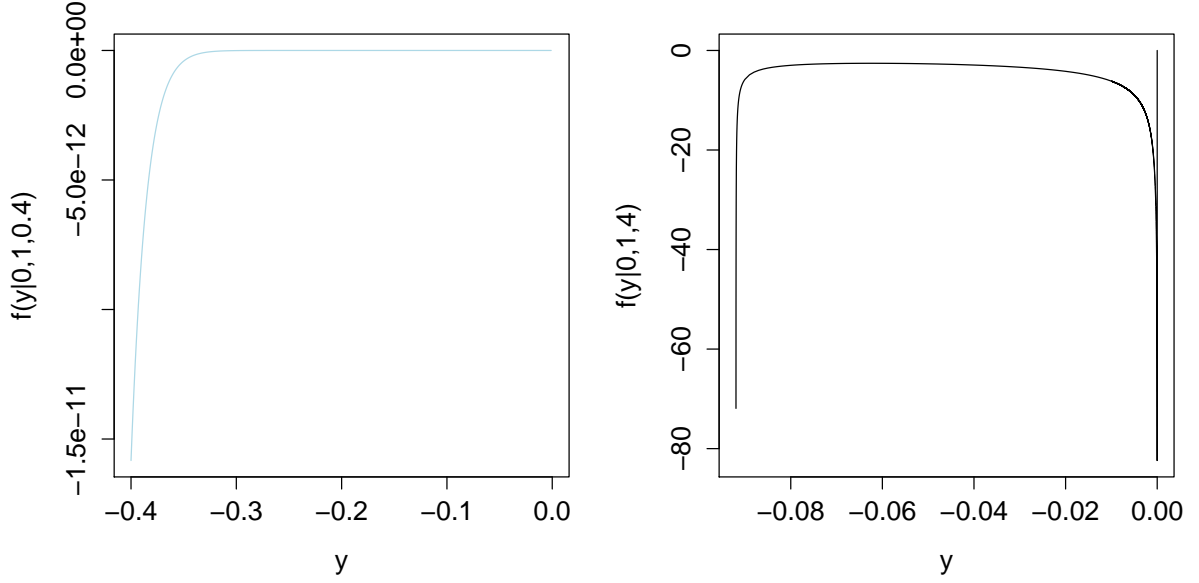
$$\frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(W_{-1}(\gamma y))^2}{2\gamma^2}\right) \frac{\exp(-W_{-1}(\gamma y))}{1 + W_{-1}(\gamma y)}. \quad (15)$$

Eelmises alapunktis analüüsisime ekstreemumpunktide olemasolu peaharu korral. Lahendini (14) jõudmisel ei olnud vaja teha harulisi eristusi. Et sama analüüsi korrata kõrvalharu jaoks, piisab vaid kontrollida, milliste γ väärtuste korral kehtib $W(\gamma y) < -1$. Võrratuse

$$\frac{-1 - \gamma^2 \pm \sqrt{\gamma^4 - 6\gamma^2 + 1}}{2} < -1 \Leftrightarrow \gamma^2 - 1 > \pm \sqrt{\gamma^4 - 6\gamma^2 + 1}$$

lahendamisel positiivsete parameetrite γ jaoks saame $\gamma \geq \sqrt{2} + 1$. Seega Lambert'i $W \times N(0, 1)$ tihedusfunktsiooni kõrvalharu komponendil on kaks ekstreemumpunkti juhul, kui asümmeetriaparameter $\gamma \in (\sqrt{2} + 1, \infty)$ (vt joonis 5, paremal). Vaatame võrduse (12) vasakut poolt kõrvalharu korral juhul kui $\gamma \in (0, \sqrt{2} + 1)$. Kuna siis võrrandil (13) lahendid puuduvad ning selle vasak pool on positiivne ja nimetajas tegur $(1 + W(\gamma y))^3$ negatiivne, siis on tuletise avaldis positiivne. Seega, kui $\gamma \in (0, \sqrt{2} + 1)$, siis on tihedusfunktsiooni graafik monotoonselt kasvav piirkonnas $y \in \left(-\frac{1}{\gamma_e}, 0\right)$ (vt joonis 5, vasakul).

Näeme, et tihedusfunktsiooni kõrvalharu komponent hakkab kahanema, kui y läheneb punktile $-\frac{1}{\gamma_e}$ paremalt. Vaatame, mis selles protsessis juhtub. On selge, et $W_{-1}(\gamma y) \rightarrow -1$, $(W_{-1}(\gamma y))^2 \rightarrow 1$ ja $1 + W_{-1}(\gamma y)$ läheb nulli vasakult. Neid kasutades saame $\frac{\exp(-W_{-1}(\gamma y))}{1 + W_{-1}(\gamma y)} \rightarrow -\infty$. Sellest tulenevalt protsessis $y \rightarrow -\frac{1}{\gamma_e} +$ koondub kõrvalharu komponent (15) miinus lõpmatusse.



Joonis 5: Lambert'i $W \times N(0,1)$ tihedusfunktsiooni kõrvalharu komponent asümmeetria-
parameetri $\gamma = 0.4$ ja $\gamma = 4$ korral

2.3.3 Lambert'i $W \times N(0,1)$ jaotus- ja tihedusfunktsioon

Vaatame nüüd mõlemat haru komponenti kasutades Lambert'i $W \times N(0,1)$ jaotus- ja tihedusfunktsiooni kuju. Valemi (7) põhjal saame juhusliku suuruse Y jaotusfunktsiooniks

$$F_Y(y|0,1,\gamma) = \begin{cases} 0, & \text{kui } y \leq -\frac{1}{\gamma e}; \\ \Phi\left(\frac{W_0(\gamma y)}{\gamma}\right) - \Phi\left(\frac{W_{-1}(\gamma y)}{\gamma}\right), & \text{kui } -\frac{1}{\gamma e} < y < 0; \\ \Phi\left(\frac{W_0(\gamma y)}{\gamma}\right), & \text{kui } y \geq 0. \end{cases}$$

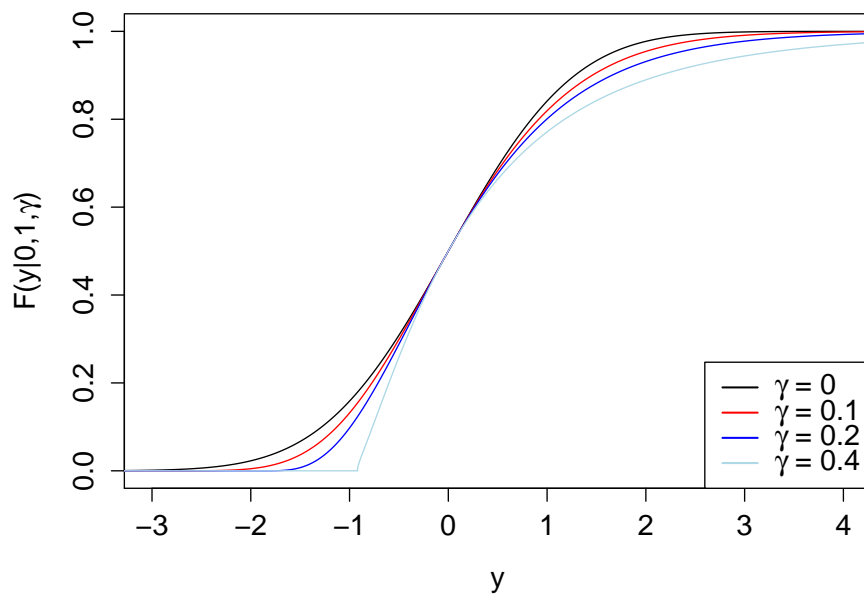
Valemi (8) abil saame tihedusfunktsiooniks

$$f_Y(y|0,1,\gamma) = \begin{cases} 0, & \text{kui } y \leq -\frac{1}{\gamma e}; \\ \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(W_0(\gamma y))^2}{2\gamma^2}\right) \frac{\exp(-W_0(\gamma y))}{1+W_0(\gamma y)} \\ \quad - \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(W_{-1}(\gamma y))^2}{2\gamma^2}\right) \frac{\exp(-W_{-1}(\gamma y))}{1+W_{-1}(\gamma y)}, & \text{kui } -\frac{1}{\gamma e} < y < 0; \\ \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(W_0(\gamma y))^2}{2\gamma^2}\right) \frac{\exp(-W_0(\gamma y))}{1+W_0(\gamma y)}, & \text{kui } y \geq 0. \end{cases} \quad (16)$$

Tihedusfunktsioon kasvab, kui $y \rightarrow -\frac{1}{\gamma e} +$. Nagu eelnevalt mainitud, siis selles protsessis peaharu komponent kasvab ning kõrvalharu komponent kahaneb, seega pea- ja kõrvalharu vastavalt omavahel lahutades, kasvab tihedusfunktsioon väga kiiresti. See kasv on tingitud

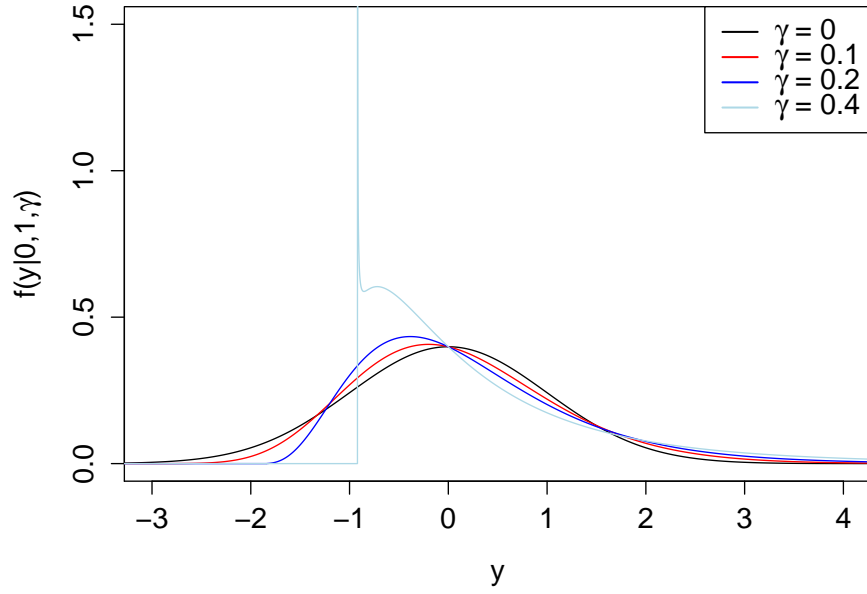
Lambert'i W funktsiooni määramispiirkonna kitsendusest, millest tulenevalt lõigatakse ära jaotuse vasak saba, kus $y \leq -\frac{1}{\gamma e}$. Nende välja jäänud väärtuste tõenäosusmass viiakse üle punkti parempoolsesse ümbrusesse. Mida suurem on asümmeetriaparaameeter, seda rohkem Lambert'i W teisendus maha löikab.

Vaatame kõigepealt saadud jaotus- ja tihedusfunktsiooni graafikut asümmeetriaparaameetrite $\gamma \in \{0, 0.1, 0.2, 0.4\}$ korral (vt joonis 6, joonis 7). Näeme, et selliste asümmeetriaparaameetrite korral saame kahe ekstreemumpunktiga tihedusfunktsiooni. Sarnaselt juhuga, kus käsitlesime ainult peaharu komponendi väärtusi, tuleb $\gamma \in \{0.1, 0.2\}$ korral vaadata funktsiooni punkti $-\frac{1}{\gamma e}$ paremas ümbruses lähemalt teise ekstreemumpunkti eristamiseks. Alapunktis, kus uurisime tihedusfunktsiooni peaharu komponenti, nägime, et sellel on kaks ekstreemumpunkti, kui $\gamma \in (0, \sqrt{2} - 1)$. Samas asümmeetriaparaameetri vahemikus oli kõrvalharu komponent aga negatiivne ning monotoonselt kasvav. Seega on mõlema haru väärtusi kasutades Lambert'i $W \times N(0, 1)$ tihedusfunktsioon samuti kahe ekstreemumpunktiga.

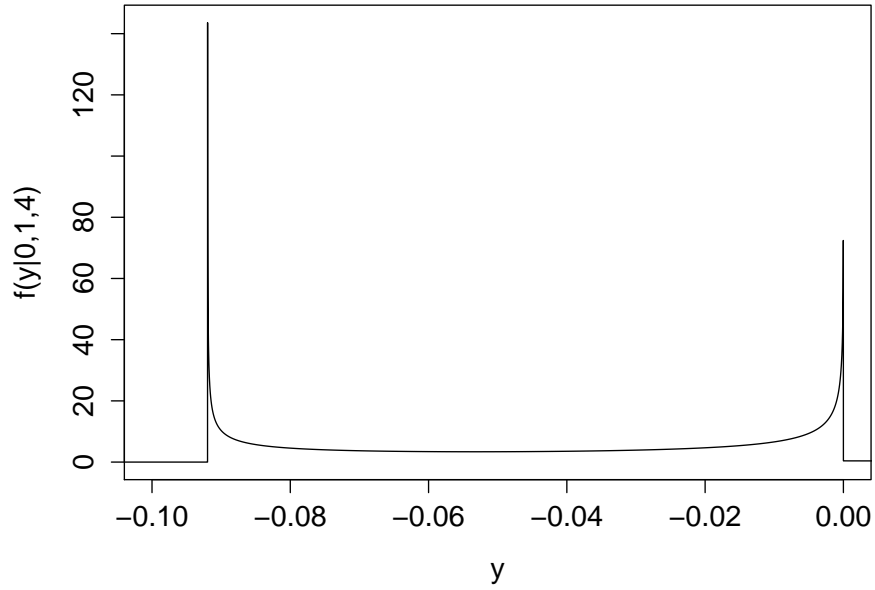


Joonis 6: Lambert'i $W \times N(0, 1)$ jaotusfunktsioon erinevate asümmeetriaparaameetri väärtuste korral

Vaatame funktsiooni ka suurema asümmeetriaparaameetri korral (vt joonis 8 $\gamma = 4$ jaoks). Meenutame, et vahemikus $\gamma > \sqrt{2} + 1$ on peaharu komponent monotoonselt kahanev ning kõrvalharu komponent kahe ekstreemumpunktiga. Seega selles piirkonnas on kõrvalharu komponendi kujust tulenevalt tihedusfunktsioon samuti kahe ekstreemumpunktiga.



Joonis 7: Lambert'i $W \times N(0, 1)$ tihedusfunktsioon erinevate asümmeetriaparametri väärtuste korral



Joonis 8: Lambert'i $W \times N(0, 1)$ tihedusfunktsioon asümmeetriaparametri $\gamma = 4$ korral

Vahemikus $\gamma \in (\sqrt{2} - 1, \sqrt{2} + 1)$ on peaharu komponent monotoonselt kahanev ja kõrvalharu komponent negatiivne ning monotoonselt kasvav. Seega sellise asümmeetria korral on tihedusfunktsioon ühe sabaga ja monotoonselt kahanev.

Lõpetuseks uurime, kui suure osa tihedusest katab peaharu komponent ja kõrvalharu

komponent. Peaharu komponendi korral saame

$$\begin{aligned}
& \int_{-\frac{1}{\gamma e}}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(W_0(\gamma x))^2}{2\gamma^2}\right) \frac{\exp(-W_0(\gamma x))}{1+W_0(\gamma x)} dx \\
&= \frac{1}{2} \int_{-\frac{1}{\gamma e}}^{\infty} \frac{2}{\sqrt{\pi}} \exp\left(-\left(\frac{W_0(\gamma x)}{\sqrt{2}\gamma}\right)^2\right) d\left(\frac{W_0(\gamma x)}{\sqrt{2}\gamma}\right) \\
&= \frac{1}{2} \left(1 - \operatorname{erf}\left(-\frac{1}{\sqrt{2}\gamma}\right)\right) = \frac{1}{2} + \frac{1}{2} \operatorname{erf}\left(\frac{1}{\sqrt{2}\gamma}\right).
\end{aligned} \tag{17}$$

Saime γ -st sõltuva funktsiooni. Positiivsete asümmeetriaparaameetrite väärtuste korral jääb saadud suuruse väärtus lõiku $[0.5, 1]$. Mida suurem on γ väärtus, seda suuremat rolli hakkab mängima kõrvalharu komponent Lambert'i $W \times N(0, 1)$ tiheduse arvutamisel. Analoogiliselt saame üle kõrvalharu määramispiirkonna integreerides

$$\begin{aligned}
& \int_0^{-\frac{1}{\gamma e}} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(W_{-1}(\gamma x))^2}{2\gamma^2}\right) \frac{\exp(-W_{-1}(\gamma x))}{1+W_{-1}(\gamma x)} dx \\
&= \frac{1}{2} \int_0^{-\frac{1}{\gamma e}} \frac{2}{\sqrt{\pi}} \exp\left(-\left(\frac{W_{-1}(\gamma x)}{\sqrt{2}\gamma}\right)^2\right) d\left(\frac{W_{-1}(\gamma x)}{\sqrt{2}\gamma}\right) \\
&= \frac{1}{2} \left(\operatorname{erf}\left(\frac{1}{\sqrt{2}\gamma}\right) - 1\right) = \frac{1}{2} \operatorname{erf}\left(\frac{1}{\sqrt{2}\gamma}\right) - \frac{1}{2}.
\end{aligned} \tag{18}$$

Selle suuruse väärtus jääb lõiku $[-0.5, 0]$. Lambert'i $W \times N(0, 1)$ tiheduse arvutamise valemit (16) järgides, saame integraalist (17) integraali (18) lahutades kontrolliks leida tihedusfunktsiooni aluse pindala, mis on 1.

2.3.4 Võrdlus asümmeetrilise normaaljaotusega

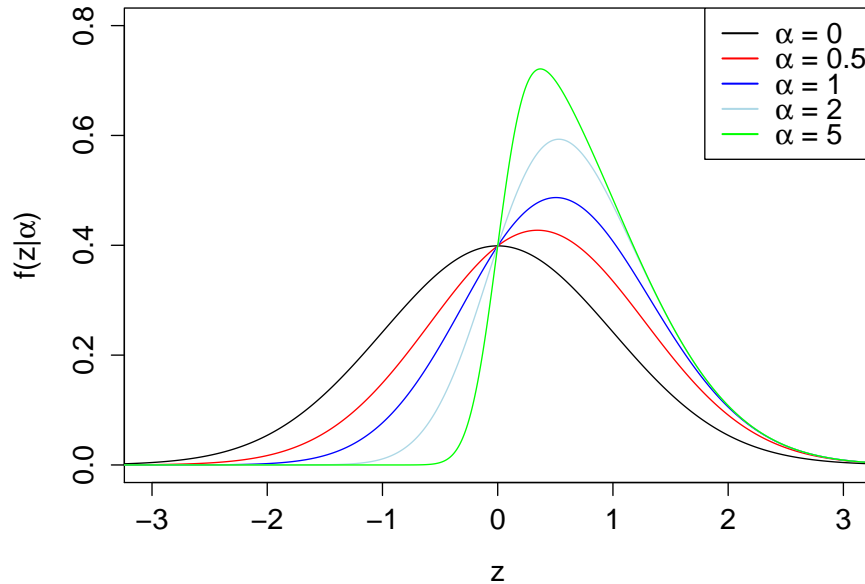
Asümmeetriliste andmete analüüsis on senini väga laialdaselt kasutusel asümmeetriline normaaljaotus. Viimaste aastakümnete jooksul on seda saatnud suur edu ning uuritud väga põhjalikult. Toome siinkohal jaotuse definitsiooni ning uurime selle momente ja asümmeetriakordajad. Asümmeetrilise normaaljaotuse kohta saab pikemalt lugeda näiteks Azzalini (1985) ja Azzalini (2005) artiklitest.

Definitsioon 5. (Azzalini (1985)) Kui juhusliku suuruse Z tihedusfunktsioon avaldub kujul

$$f(z|\alpha) = 2\phi(z)\Phi(\alpha z), \quad z \in \mathbb{R}, \tag{19}$$

kus ϕ on standardse normaaljaotuse tihedusfunktsioon ning Φ standardse normaaljaotuse jaotusfunktsioon, siis ütleme, et Z on asümmeetrilise normaaljaotusega paraameetriga α .

Parameeter α võib omandada kõiki väärtusi reaalteljel. Kui $\alpha = 0$, siis tihedusfunktsioon (19) taandub standardse normaaljaotuse tihedusfunktsiooniks. Positiivse asümmeetriapara-meetri α korral on jaotus paremale kaldu (vt joonis 9), negatiivse korral vasakule.



Joonis 9: Asümmeetrilise normaaljaotuse tihedusfunktsioon erinevate asümmeetriapara-meetri väärtuste korral

Töö praktilises osas läheb vaja ka asukoha-skaala asümmeetrilist normaaljaotust, mida tähistame $SN(\xi, \omega, \alpha)$, kus lisaks asümmeetriapara-meetrile α on asukohapara-meeter $\xi \in \mathbb{R}$ ja skaalapara-meeter $\omega > 0$. Selle definitsioonini jõuab, kasutades fakti, et kui $Z \sim SN(\alpha)$, siis $\xi + \omega Z \sim SN(\xi, \omega, \alpha)$ (Azzalini (2005)). Seega asendame definitsioonis 5 z asemele $\frac{z - \xi}{\omega}$ ning skaleerime tihedust vastavalt.

Definitsioon 6. Kui juhusliku suuruse Z tihedusfunktsioon avaldub kujul

$$f(z|\xi, \omega, \alpha) = \frac{2}{\omega} \phi\left(\frac{z - \xi}{\omega}\right) \Phi\left(\alpha \left(\frac{z - \xi}{\omega}\right)\right), \quad z \in \mathbb{R}, \omega > 0,$$

kus ϕ on standardse normaaljaotuse tihedusfunktsioon ning Φ standardse normaaljaotuse jaotusfunktsioon, siis ütleme, et Z on asukoha-skaala asümmeetrilise normaaljaotusega juhuslik suurus asukohapara-meetriga ξ , skaalapara-meetriga ω ja asümmeetriapara-meetriga α , ning kirjutame $Z \sim SN(\xi, \omega, \alpha)$.

Olgu Z asümmeetrilise standardse normaaljaotusega juhuslik suurus. Siis Azzalini (1985)

kohaselt on selle momente genereeriv funktsioon kujul

$$M_Z(t) = 2 \exp\left(\frac{t^2}{2}\right) \Phi(\delta t),$$

kus $\delta = \frac{\alpha}{\sqrt{1+\alpha^2}}$. Seda kasutades leiame esimesed kolm (tsentreerimata) momenti. Esimese momendi jaoks leiame

$$M'_Z(t) = 2t \exp\left(\frac{t^2}{2}\right) \Phi(\delta t) + 2 \exp\left(\frac{t^2}{2}\right) \delta \phi(\delta t),$$

millest

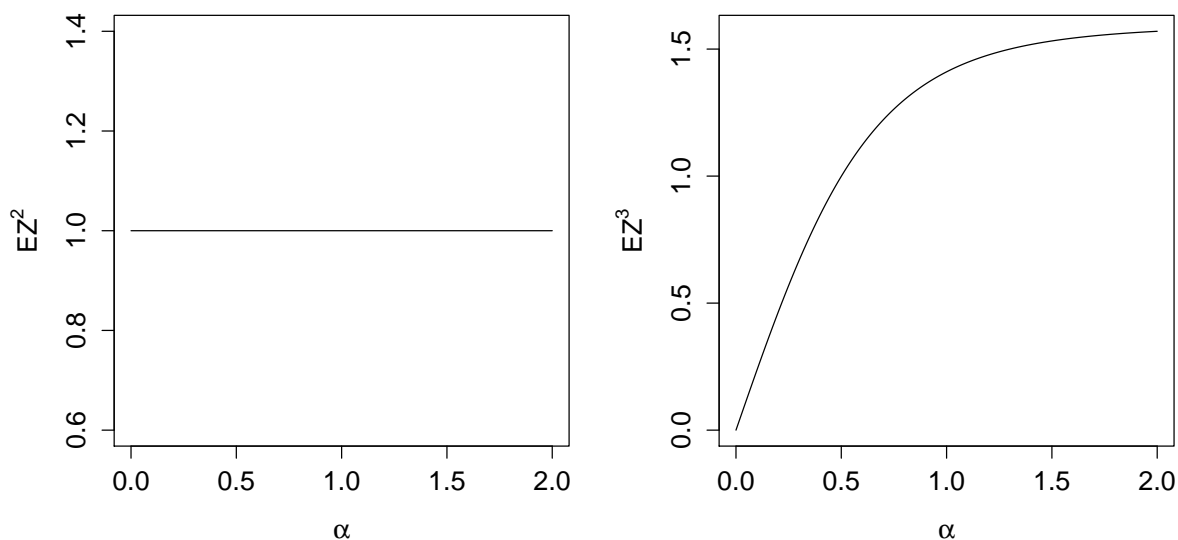
$$EZ = M'_Z(0) = 2\delta \frac{1}{\sqrt{2\pi}} = \delta \sqrt{\frac{2}{\pi}}.$$

Analoogiliselt jätkates leiame

$$EZ^2 = M''_Z(0) = 1$$

ja

$$EZ^3 = M'''_Z(0) = (3 - \delta^2)\delta \sqrt{\frac{2}{\pi}}.$$



Joonis 10: Asümmeetrilise normaaljaotuse teist ja kolmandat järku momendid

Saame leida asümmeetriakordaja $\gamma_1(Z) = \frac{E(Z - \mu_Z)^3}{\sigma_Z^3} = \frac{EZ^3 - 3\mu_Z\sigma_Z^2 - \mu_Z^3}{\sigma_Z^3}$ väärtuse

$$\gamma_1(\alpha) = \frac{4 - \pi}{2} \operatorname{sgn}(\alpha) \left(\frac{\frac{2\delta^2}{\pi}}{1 - \frac{2\delta^2}{\pi}} \right)^{\frac{3}{2}},$$

kus

$$\operatorname{sgn}(\alpha) = \begin{cases} -1, & \text{kui } \alpha < 0; \\ 0, & \text{kui } \alpha = 0; \\ 1, & \text{kui } \alpha > 0. \end{cases}$$

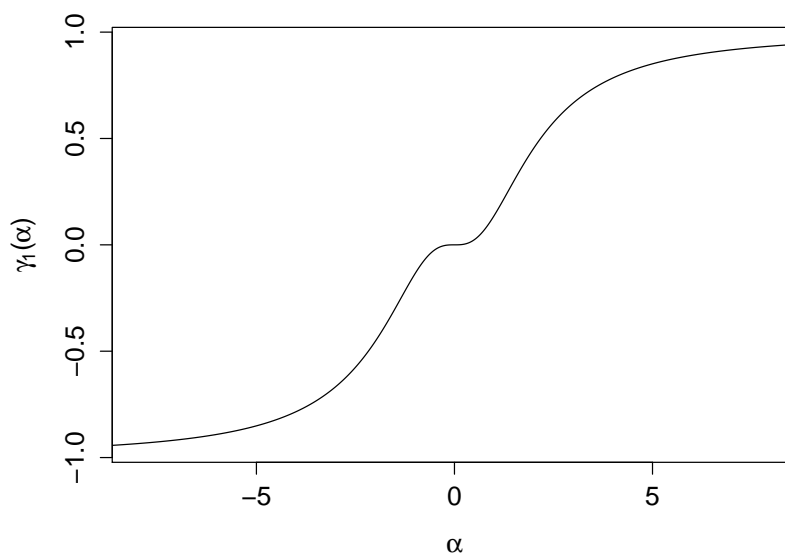
Uurime, mis vahemikus saavad $\gamma_1(\alpha)$ väärtused muutuda. Selleks leiame piirväärtused

$\lim_{\alpha \rightarrow \infty} \gamma_1(\alpha)$ ja $\lim_{\alpha \rightarrow -\infty} \gamma_1(\alpha)$. Kõigepealt paneme tähele, et

$$\lim_{\alpha \rightarrow \pm\infty} \delta^2 = \frac{\alpha^2}{1 + \alpha^2} = \frac{1}{\frac{1}{\alpha^2} + 1} = 1,$$

millest

$$\begin{aligned} \lim_{\alpha \rightarrow \pm\infty} \left(\frac{\frac{2\delta^2}{\pi}}{1 - \frac{2\delta^2}{\pi}} \right)^{\frac{3}{2}} &= \lim_{\alpha \rightarrow \pm\infty} \left(\frac{1}{\frac{\pi}{2\delta^2} - 1} \right)^{\frac{3}{2}} \\ &= \left(\frac{1}{\frac{\pi}{2} - 1} \right)^{\frac{3}{2}} = \frac{2^{\frac{3}{2}}}{(\pi - 2)^{\frac{3}{2}}} \approx 2.3189. \end{aligned}$$



Joonis 11: Asümmeetrilise normaaljaotuse asümmeetriakordaja väärtused parameetri α muutudes

Saame nüüd otsitavateks piirväärtusteks

$$\lim_{\alpha \rightarrow \infty} \gamma_1(\alpha) = \frac{4 - \pi}{2} \frac{2^{\frac{3}{2}}}{(\pi - 2)^{\frac{3}{2}}} \approx 0.9953$$

ja

$$\lim_{\alpha \rightarrow -\infty} \gamma_1(\alpha) = -\frac{4-\pi}{2} \frac{2^{\frac{3}{2}}}{(\pi-2)^{\frac{3}{2}}} \approx -0.9953.$$

Saime, et asümmeetrilise normaaljaotuse asümmeetriakordaja on piiratud vahemikku $(-0.9953, 0.9953)$. Joonisel 11 on näha, et asümmeetriakordaja piiride lähedale jõutakse juba umbes jaotuse parameetri $\alpha = \pm 8$ korral. Absoluutväärtuselt suuremate α väärtuste korral jääb $\gamma_1(\alpha) \approx 0.9953$.

Analüüsime nüüd asümmeetriakordaja käitumist normaaljaotusega Lambert'i W juhuslike suuruste korral. Olgu X normaaljaotusega juhuslik suurus ning Y vastav Lambert'i W juhuslik suurus. Goerg (2011) toob välja valemi selle momentide arvutamiseks

$$E(Y^n) = \frac{1}{n^n} \frac{\partial^n}{\partial \gamma^n} M_X(\gamma n) = \frac{1}{n^n} \frac{\partial^n}{\partial \gamma^n} \exp \left(\gamma n \mu_X + \frac{\gamma^2 n^2 \sigma_X^2}{2} \right).$$

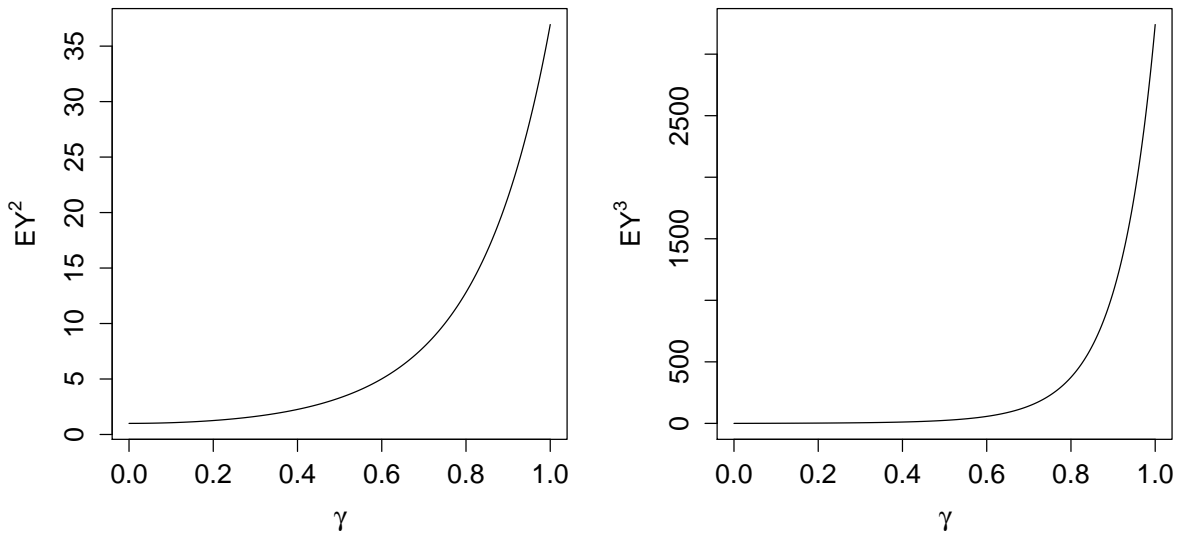
Seda valemit kasutades leiame Lambert'i $W \times N(0, 1)$ jaotuse momendid kolmanda jär-
guni:

$$E(Y) = \frac{1}{1^1} \frac{\partial}{\partial \gamma} \exp \left(\frac{\gamma^2 1^2}{2} \right) = \gamma \exp \left(\frac{\gamma^2}{2} \right),$$

$$E(Y^2) = \frac{1}{2^2} \frac{\partial^2}{\partial \gamma^2} \exp \left(\frac{\gamma^2 2^2}{2} \right) = (1 + 4\gamma^2) \exp(2\gamma^2)$$

ja

$$E(Y^3) = \frac{1}{3^3} \frac{\partial^3}{\partial \gamma^3} \exp \left(\frac{\gamma^2 3^2}{2} \right) = 9\gamma(1 + 3\gamma^2) \exp \left(\frac{9\gamma^2}{2} \right).$$



Joonis 12: Lambert'i $W \times N(0, 1)$ teist ja kolmandat järku momendid

Saadud momentide avaldiste põhjal saab tuletada juhusliku suuruse Y asümmeetriakordaja $\gamma_1(Y) = \frac{E(Y - \mu_Y)^3}{\sigma_Y^3} = \frac{EY^3 - 3\mu_Y\sigma_Y^2 - \mu_Y^3}{\sigma_Y^3}$ väärtuse

$$\gamma_1(\gamma) = \gamma \left(\frac{e^{3\gamma^2}(9 + 27\gamma^2) - e^{\gamma^2}(3 + 12\gamma^2) + 2\gamma^2}{(e^{\gamma^2}(1 + 4\gamma^2) - \gamma^2)^{\frac{3}{2}}} \right).$$

Asümmeetriakordaja väärtused pole tõkestatud, sest vähendades lugejat ja suurendades nimetajat saame kasutada hinnangut

$$\begin{aligned} \frac{e^{3\gamma^2}(9 + 27\gamma^2) - e^{\gamma^2}(3 + 12\gamma^2) + 2\gamma^2}{(e^{\gamma^2}(1 + 4\gamma^2) - \gamma^2)^{\frac{3}{2}}} &\geq \frac{e^{3\gamma^2}(9 + 27\gamma^2) - e^{\gamma^2}(3 + 12\gamma^2)}{e^{\frac{3}{2}\gamma^2}(1 + 4\gamma^2)^{\frac{3}{2}}} \\ &= e^{\frac{3}{2}\gamma^2} \frac{9 + 27\gamma^2}{(1 + 4\gamma^2)^{\frac{3}{2}}} - e^{-\frac{\gamma^2}{2}} \frac{3 + 12\gamma^2}{(1 + 4\gamma^2)^{\frac{3}{2}}}. \end{aligned}$$

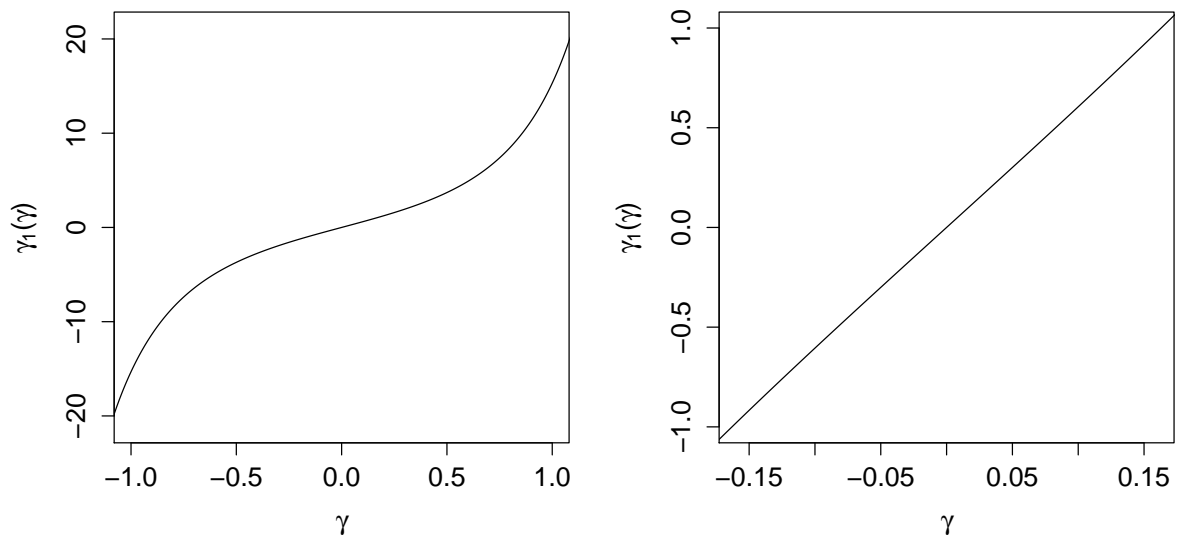
Eksponentfunktsioon kasvab kiiremini kui astmefunktsioon, seetõttu kui $\gamma \rightarrow \pm\infty$, siis läheb esimene liige lõpmatusse ja teine koondub nulli. Sellest järeldub, et

$$\lim_{\gamma \rightarrow \pm\infty} \gamma_1(\gamma) = \pm\infty.$$

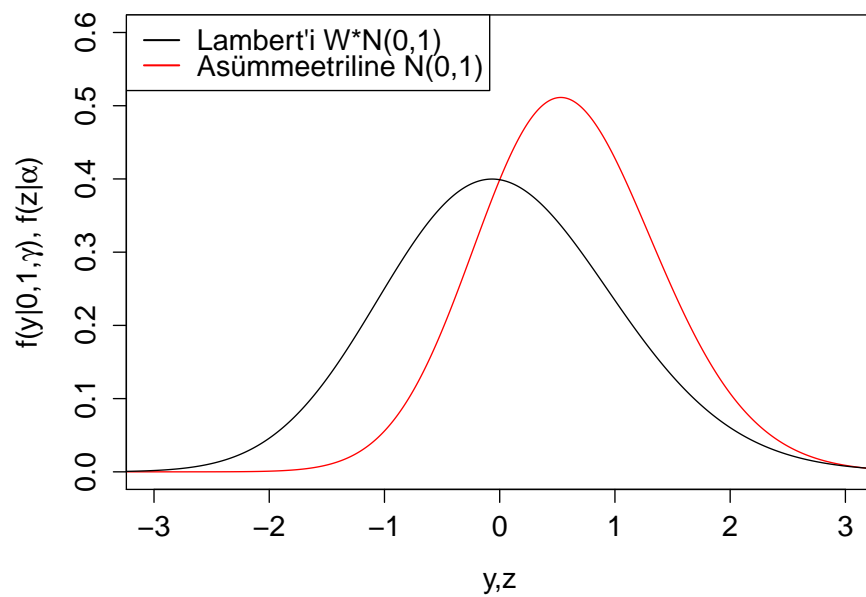
Meenutades, et asümmeetrilise normaaljaotuse asümmeetriakordaja katab vaid vahemikku $(-0.9953, 0.9953)$, siis on selge, et Lambert'i $W \times N(0, 1)$ jaotus on palju mitmekülgsem ning võimaldab kirjeldada ka ekstreemsema asümmeetrilisusega andmeid.

Joonisel 13 vasakul on näha, et normaaljaotusega Lambert'i W jaotuse asümmeetriakordaja varieerub juba ainult parameetri $\gamma \in (-1, 1)$ väärtuste korral väga laias vahemikus, $\gamma_1(\gamma) \in (-15, 15)$. Kui $|\gamma| > 1$, siis asümmeetriakordaja absoluutväärtus kasvab veelgi kiiremini. Vaadates asümmeetriakordajat lähemalt samas muutumiskiirkonnas kui asümmeetrilise normaaljaotuse korral, näeme, et graafik meenutab tõusvat sirget. Sama muutumiskiirkonna katavad Lambert'i $W \times N(0, 1)$ juhuslikud suurused juba parameetri $\gamma \in (-0.163, 0.163)$ korral (vt joonis 13, paremal).

Lõpetuseks visualiseerime asümmeetrilise normaaljaotuse ja Lambert'i $W \times N(0, 1)$ jaotuse tihedusfunktsioone sama asümmeetriakordaja γ_1 korral (vt joonis 14). Valime selleks näiteks $\gamma_1 = 0.2$. Sellisel juhul on asümmeetrilise normaaljaotuse parameeter $\alpha \approx 1.1988$ ja Lambert'i W korral $\gamma \approx 0.0332$. Näeme, et tihedusfunktsioonide kujud erinevad üksteisest olulisel määral. Jaotuste tipud on erineva kõrguse ja asukohaga, sabasid on antud joonise põhjal raske võrrelda.



Joonis 13: Lambert'i $W \times N(0, 1)$ asümmeetriakordaja väärtused parameetri $\gamma \in (-1, 1)$ ja $\gamma \in (-0.163, 0.163)$ korral



Joonis 14: Asümmeetrilise normaaljaotuse ja Lambert'i $W \times N(0, 1)$ tihedusfunktsioonid asümmeetriakordaja $\gamma_1 = 0.2$ korral

Lisas B.2 leiab Rööp (2015) magistritöö põhjal koostatud animeeritava graafiku koodi, kus on peal mõlema jaotuse tihedused.

2.4 Lambert'i W t-jaotus

T-jaotus on üsna sarnane normaaljaotusega. Selle tihedusfunktsioon on samuti sümmeetriline ning kellakujuline, küll on sellel aga raskemad sabad kui normaaljaotusel. Tähistame t-jaotust $T(c, s, \nu)$, kus c on asukohaparameter, s skaalaparameter ning ν on vabadusastmete arv. Seega parameetervektor $\beta = (c, s, \nu)$. Uurime siinkohal tsentraalse t-jaotuse juhtumit, kus sisendjaotuses $c = 0$ ja $s = 1$ ning vaatame vastavat Lambert'i $W \times T(0, 1, \nu)$ jaotus- ja tihedusfunktsiooni. Olgu $X \sim T(0, 1, \nu)$, siis t-jaotuse omadustest teame, et $\mu_X = c = 0$ ja $\sigma_X = s\sqrt{\frac{\nu}{\nu-2}} = \sqrt{\frac{\nu}{\nu-2}}$, kui $\nu > 2$. Olgu Y Lambert'i $W \times T(0, 1, \nu)$ jaotusest juhuslik suurus.

Kasutades valemeid (7) ja (8), kus võtame vastavalt $\mu_X = 0$, saame leida Y tihedusfunktsiooni

$$f_Y(y|\theta) = \begin{cases} 0, & \text{kui } y \leq -\frac{\sigma_X}{\gamma e}; \\ \frac{\Gamma(\frac{\nu+1}{2})}{\sqrt{\nu\pi}\Gamma(\frac{\nu}{2})} \left(1 + \frac{(W_0(\gamma z)\sigma_X)^2}{\gamma^2\nu}\right)^{-\frac{\nu+1}{2}} \frac{\exp(-W_0(\gamma z))}{1+W_0(\gamma z)} - \frac{\Gamma(\frac{\nu+1}{2})}{\sqrt{\nu\pi}\Gamma(\frac{\nu}{2})} \left(1 + \frac{(W_{-1}(\gamma z)\sigma_X)^2}{\gamma^2\nu}\right)^{-\frac{\nu+1}{2}} \frac{\exp(-W_{-1}(\gamma z))}{1+W_{-1}(\gamma z)}, & \text{kui } -\frac{\sigma_X}{\gamma e} < y < 0; \\ \frac{\Gamma(\frac{\nu+1}{2})}{\sqrt{\nu\pi}\Gamma(\frac{\nu}{2})} \left(1 + \frac{(W_0(\gamma z)\sigma_X)^2}{\gamma^2\nu}\right)^{-\frac{\nu+1}{2}} \frac{\exp(-W_0(\gamma z))}{1+W_0(\gamma z)}, & \text{kui } y \geq 0, \end{cases}$$

millest saame leida jaotusfunktsiooni

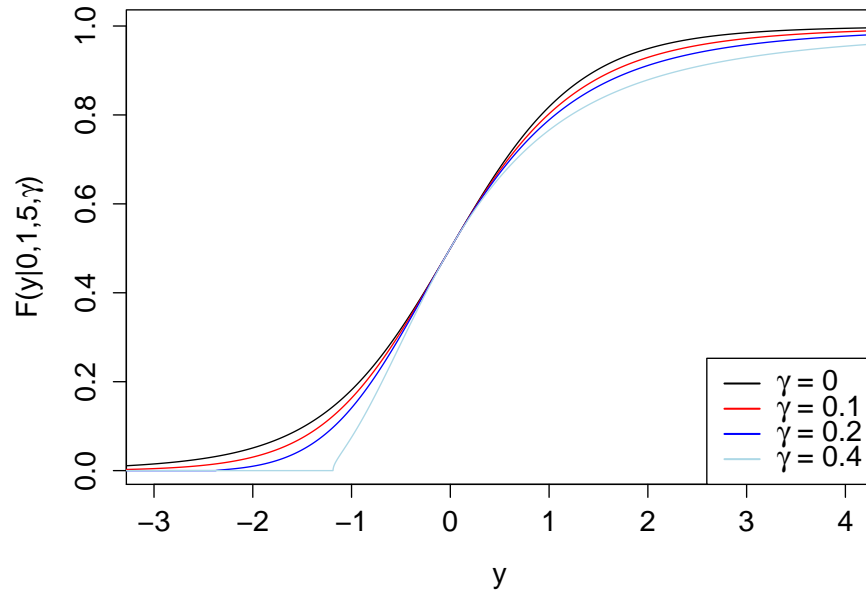
$$F_Y(y|\theta) = \begin{cases} 0, & \text{kui } y \leq -\frac{\sigma_X}{\gamma e}; \\ \frac{\Gamma(\frac{\nu+1}{2})}{\sqrt{\nu\pi}\Gamma(\frac{\nu}{2})} \int_{-\frac{1}{\gamma e}}^z \left(1 + \frac{(W_0(\gamma t)\sigma_X)^2}{\gamma^2\nu}\right)^{-\frac{\nu+1}{2}} d\left(\frac{W_0(\gamma t)\sigma_X}{\gamma}\right) - \frac{\Gamma(\frac{\nu+1}{2})}{\sqrt{\nu\pi}\Gamma(\frac{\nu}{2})} \int_{-\frac{1}{\gamma e}}^z \left(1 + \frac{(W_{-1}(\gamma t)\sigma_X)^2}{\gamma^2\nu}\right)^{-\frac{\nu+1}{2}} d\left(\frac{W_{-1}(\gamma t)\sigma_X}{\gamma}\right), & \text{kui } -\frac{\sigma_X}{\gamma e} < y < 0; \\ \frac{\Gamma(\frac{\nu+1}{2})}{\sqrt{\nu\pi}\Gamma(\frac{\nu}{2})} \int_{-\frac{1}{\gamma e}}^z \left(1 + \frac{(W_0(\gamma t)\sigma_X)^2}{\gamma^2\nu}\right)^{-\frac{\nu+1}{2}} d\left(\frac{W_0(\gamma t)\sigma_X}{\gamma}\right), & \text{kui } y \geq 0, \end{cases}$$

kus $\theta = (\beta, \gamma) = (0, 1, \nu, \gamma)$, $\sigma_X = \sqrt{\frac{\nu}{\nu-2}}$ ja $z = \frac{y}{\sigma_X}$.

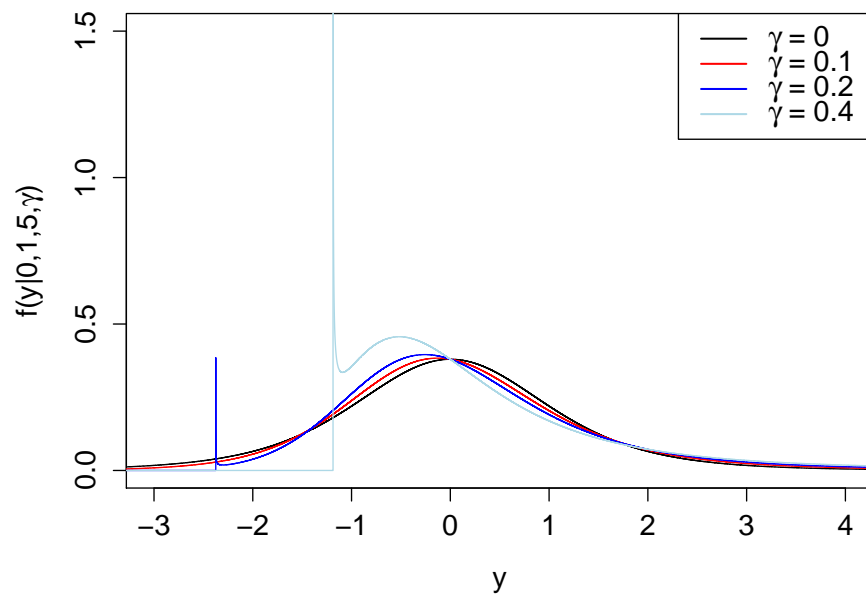
Võtame tulemuste graafiliseks esitamiseks vabadusastmete arvaks $\nu = 5$, siis $\sigma_X = \sqrt{\frac{5}{3}}$, ning asümmeetriaparametrite väärtusteks $\gamma \in \{0, 0.1, 0.2, 0.4\}$.

Joonistel 6 ja 15 toodud vastavad Lambert'i $W \times N(0, 1)$ ja Lambert'i $W \times T(0, 1, 5)$ jaotusfunktsioonide kujud asümmeetriaparametrite $\gamma \in \{0, 0.1, 0.2, 0.4\}$ korral näevad väga sarnased välja. Nimetatud jaotuste tihedusfunktsioone joonistel 7 ja 16 uurides, on näha, et t-jaotuse puhul on funktsiooni tipud madalamad nagu ka juhul kui asümmeetrisus puudub. Kui z läheneb punktile $-\frac{1}{\gamma e}$ paremalt, siis kasvab Lambert'i $W \times T(0, 1, 5)$

tihedusfunktsioon kiiremini kui sama asümmeetrilisusega Lambert'i $W \times N(0, 1)$ tihedus. Näiteks parameetri $\gamma = 0.2$ korral pidime Lambert'i $W \times N(0, 1)$ tiheduse graafikut vaatama väga lähedalt, joonisel 16 on see aga selgelt eristuv.



Joonis 15: Lambert'i $W \times T(0, 1, 5)$ jaotusfunktsioon erinevate asümmeetriaparametri väärtuste korral



Joonis 16: Lambert'i $W \times T(0, 1, 5)$ tihedusfunktsioon erinevate asümmeetriaparametri väärtuste korral

Mida väiksem on t-jaotuse vabadusastmete arv, seda raskemad on selle jaotuse sabad.

Kuna meie näites on võetud $\nu = 5$, siis põhjustab Lambert'i W rakendamine määramispiirkonna piirangu tõttu normaaljaotusega võrreldes raskema saba ära lõikamist. Äralõigatud saba tõenäosusmass koondub punkti $-\frac{1}{\gamma e}$ parempoolsesse ümbrusesse, põhjustades seal tiheduse kiirema kasvu kui Lambert'i $W \times N(0, 1)$ korral.

3 Jaotuste sobitamine andmetele

Selles alapunktis vaatame eelneva teooria praktikas kasutamist. Kuna asümmeetrilised andmed esinevad tihti kindlustusvaldkonnas, siis antud magistritöö rakendamiseks selles vallas, töötame reaalse kahjusummade andmestikuga. Nimelt sobitame kahjusummadele Lambert'i $W \times \text{Exp}(\lambda)$, Lambert'i $W \times N(\mu_X, \sigma_X)$ jaotust ning võrdluseks asukoha-skaala asümmeetrilist normaaljaotust. Uurime, milline jaotus paremini andmeid kirjeldab ning võrdleme saadud tulemusi ka teiste kindlustuses laialt levinud jaotustega. Esmalt uurime Goerg'i (2011) artikli põhjal parameetrite hindamist suurima tõepära meetodiga.

3.1 Parameetrite hindamine

Lambert'i $W \times F_X$ jaotuse andmetele sobitamiseks peame leidma parameetritele hinnangud. Kasutame selleks suurima tõepära meetodit. Eeldame, et on antud valim suurusega N sõltumatute sama jaotusega juhuslike suuruste vaatlusi $y = (y_1, y_2, \dots, y_N)$, mis pärinevad transformatsioonist (4). Seega tuleb leida hinnangud parameetervektori θ komponentidele, st sisendjaotuse parameetri(te)le β ja asümmeetriaparametrile γ .

Olgu Lambert'i $W \times F_X$ jaotusest juhusliku suuruse Y tihedusfunktsioon $f_Y(y|\beta, \gamma)$ antud valemiga (8). Siis valimi log-tõepära funktsioon on kujul

$$l(\theta|y) = \sum_{i=1}^N \ln(f_Y(y_i|\theta)).$$

Suurima tõepära hinnangud parameetritele saadakse log-tõepära maksimiseerimisel üle võimalike θ väärtuste ehk leitakse selline $\hat{\theta}_{MLE} = (\hat{\beta}_{MLE}, \hat{\gamma}_{MLE})$ nii, et

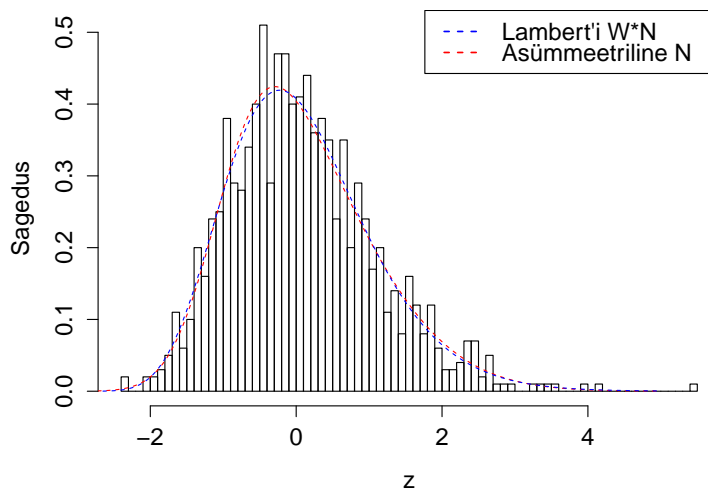
$$\hat{\theta}_{MLE} = \arg \max_{\theta} l(\theta|y).$$

Edasi arvutatakse suurima tõepära hinnangud sisendjaotuse tihedusest sõltuvalt. Üldjuhul pole kahe haru väärtusi kasutades rohkem avaldist lihtsustada võimalik. Saame mittelineaarse optimiseerimisülesande, mille lahendamiseks tuleb kasutada numbrilisi meetodeid.

3.2 Simulatsioonid

Selles alapunktis võrdleme asümmeetrilist normaaljaotust ja Lambert'i $W \times N(\mu_X, \sigma_X)$ jaotust läbi simulatsioonide. Jaotuste sobitamisel käesolevas ja järgnevas alapunktis kasutame definitsioonis 6 defineeritud asukoha-skaala asümmeetrilist normaaljaotust ning Lambert'i $W \times N(\mu_X, \sigma_X)$ jaotust. Genereerime andmeid asümmeetrilisest normaaljaotusest ning sobitame neile nii jaotust ennast kui ka Lambert'i $W \times N(\mu_X, \sigma_X)$ ning vastupidi. Uurime, kas genereeritud väärtustele ühest jaotusest võib teine alternatiiv sobida sama hästi. Sobivuse hindamiseks kasutame Kolmogorov-Smirnovi (K-S) statistikut ja Akaike informatsiooni kriteeriumit (AIC).

Genereerime esmalt asümmeetriakordajaga $\gamma = 0.1$ Lambert'i $W \times N(0, 1)$ jaotusest 1000 juhuslikku realisatsiooni. Vaatame esmalt kummagi jaotuse sobivust tiheduse histogrammilt.



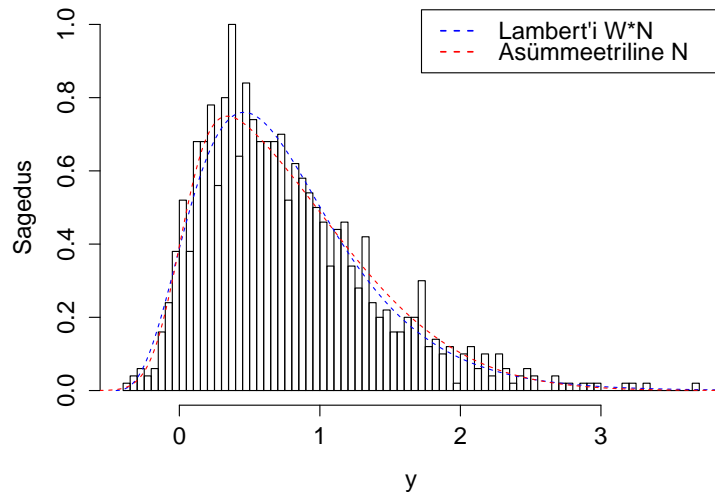
Joonis 17: Lambert'i $W \times N(0, 1)$ jaotuse realisatsioonidele sobitatud jaotused

Jaotuse tiheduse kõverad on väga sarnased, silmaga peaaegu eristamatud. Esitame kokkuvõtlikult tabeli kujul sobitamisel saadud parameetrid ja sobivuse näitajad (vt tabel 1). Näeme, et nii AIC kui Kolmogorov-Smirnovi statistiku põhjal sobituvad jaotused enam-vähem sama hästi, asümmeetriline normaaljaotus sobitub natuke kehvemini. Lambert'i $W \times N$ sobitamisel saadi üsna teoreetilise jaotuse lähedased parameetrid. Seega andmete väikese asümmeetrilisuse korral leidub erineva parametrizeeringuga sarnaseid tiheduse kujusid.

Tabel 1: Lambert'i $W \times N(0, 1)$ jaotuse juhuslikele realisatsioonidele sobitatud jaotuste hinnatud parameetrid ja jaotuste sobivuse statistikud

	Asümmeetrilise normaaljaotuse sobitamine	Lambert'i $W \times N$ sobitamine
Asukohaparameter	$\xi = -1.036428$	$\mu = -0.022900$
Skaalaparameter	$\omega = 1.502930$	$\sigma = 0.973540$
Asümmeetriaparameter	$\alpha = 2.628544$	$\gamma = 0.106280$
Kolmogorov-Smirnovi statistik	0.015494	0.012572
K-S statistiku p-väärtus	0.9700	0.9974
AIC	2781.076	2778.768

Genereerime nüüd standardsest asümmeetrilisest normaaljaotusest samuti 1000 juhuslikku väärtust. Asümmeetriaparaameetiks valime $\alpha = 6$, et andmed oleks suurema asümmeetriaga. Uurime, kui hästi siis kumbki jaotus sobitub. Vaatame mõlema jaotuse sobivust esmalt graafiliselt.



Joonis 18: Asümmeetrilise normaaljaotuse realisatsioonidele sobitatud jaotused

Sel korral on tiheduse graafikud juba selgemini eristuvad. Taas toome sobitamisel saadud parameetrite hinnangud ja sobivuse näitajad tabeli kujul (vt tabel 2). Sobitamisel näeme, et asümmeetriline normaaljaotus on vaid veidi parem kui alternatiivne jaotus, Kolmogorov-Smirnovi statistiku ja AIC väärtused on praktiliselt võrdsed.

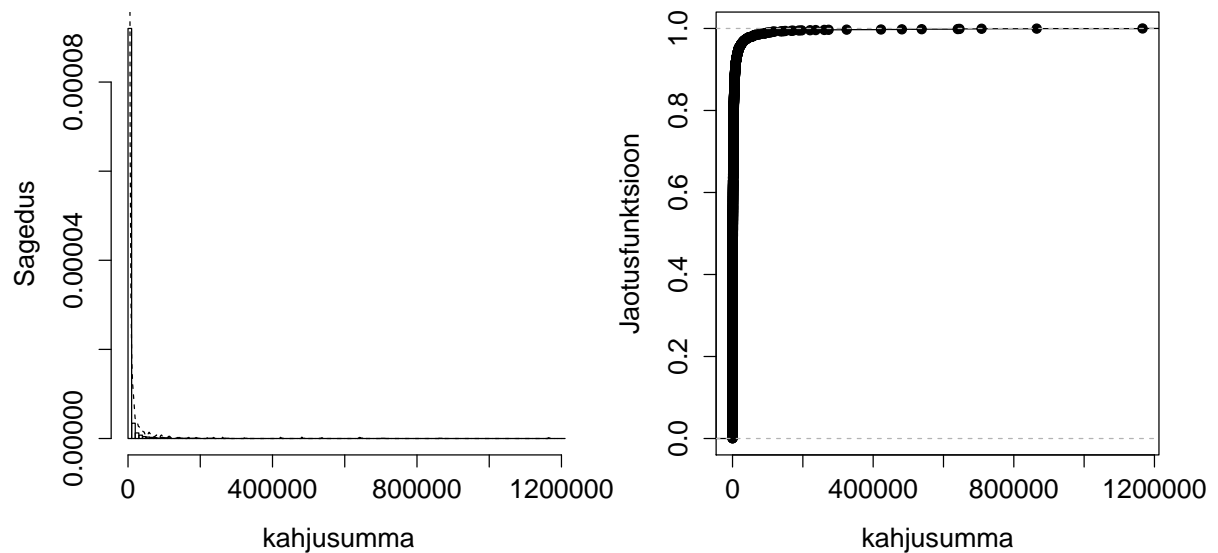
Tabel 2: Asümmeetrilise normaaljaotuse juhuslikele realisatsioonidele sobitatud jaotuste hinnatud parameetrid ja jaotuste sobivuse statistikud

	Asümmeetrilise normaaljaotuse sobitamine	Lambert'i $W \times N$ sobitamine
Asukohaparameter	$\xi = 0.006949$	$\mu = 0.662500$
Skaalaparameter	$\omega = 0.979599$	$\sigma = 0.564140$
Asümmeetriaparameter	$\alpha = 5.752658$	$\gamma = 0.185990$
Kolmogorov-Smirnovi statistik	0.020769	0.022579
K-S statistiku p-väärtus	0.7815	0.6878
AIC	1660.954	1662.574

Näeme, et teostatud simulatsioonide põhjal on jaotused väga sarnased. Sobivuse näitajad olid mõlemal juhul enam-vähem võrdsed. Jaotus, millest juhuslikud väärtused olid genereeritud, sobitus vaevumärgatavalt paremini, kui vastav alternatiiv. Ka graafiliselt on näha, et silmnähtavalt sobitusid mõlemad jaotused andmetele hästi. Lambert'i $W \times N(0, 1)$ jaotusest genereeritud väärtuste korral kehtib see ainult väiksemate asümmeetriaparaametrete γ korral, tulenevalt alapunktis 2.3.4 teostatud analüüsist, milles selgus selle tiheduse eripärane kuju. Ka asümmeetrilisest normaaljaotusest genereeritud väärtuste puhul muutus Lambert'i $W \times N$ jaotus suurema asümmeetrilisuse puhul vähem sobivamaks. Lambert'i $W \times N$ jaotus pakub hea alternatiivi traditsioonilisele asümmeetrilisele normaaljaotusele, olles väiksema asümmeetrilisuse korral üsna sarnane, kuid võimaldades kirjeldada andmeid ka suurema asümmeetriaga.

3.3 Jaotuste sobitamine kahjusummadele

Selle alapunkti eesmärk on tutvustada Lambert'i $W \times F_X$ juhuslike suuruste teooria rakendamisvõimalusi kindlustusvaldkonnas. Meil on andmestik 2 797 kahju suurusega. Minimaalne kahju on 15.27 ning maksimaalne 1 166 171 ühikut. Keskmise kahjusumma suurus on 6 671.067 ning mediaan 952.97. Kahjusummade varieeruvus on väga suur üksikute suurte (üle 200 000 ühiku suuruste) kahjude tõttu, mida on kokku 13 tükki.



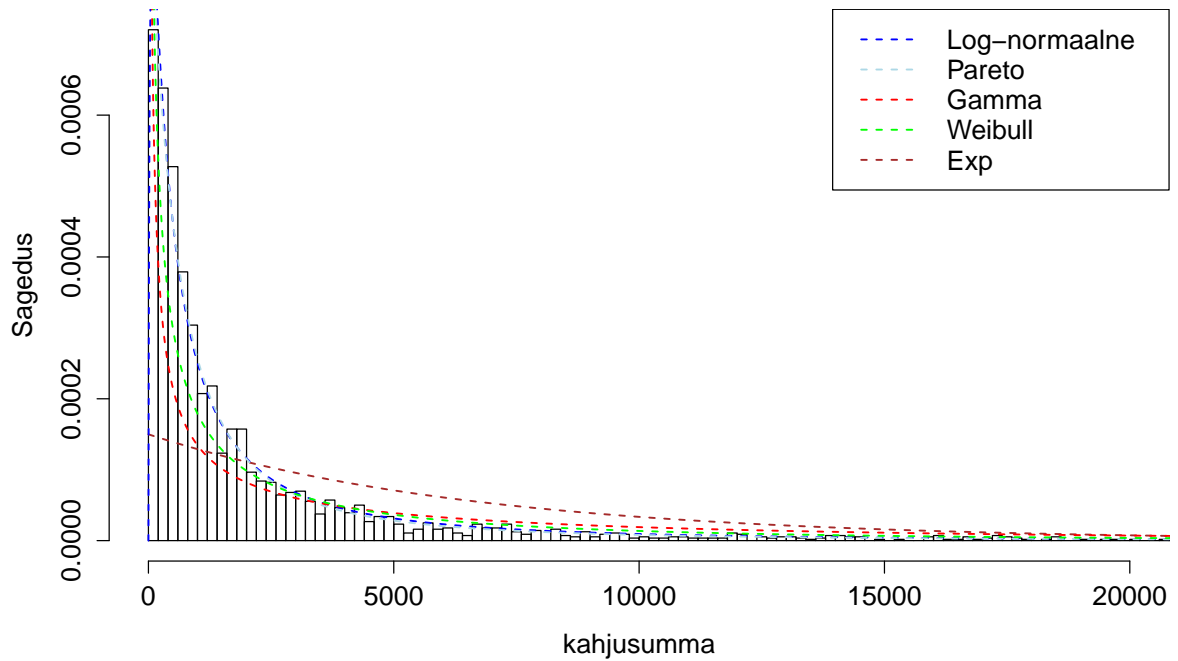
Joonis 19: Kahjusummade histogramm (tihedus) ja empiiriline jaotus

Samale andmestikule on Puksand oma 2015. aastal valminud magistritöös sobitanud levinumaid kahjujaotusi: eksponentjaotust, log-normaalset jaotust, Pareto jaotust, gamma-jaotust ja Weibulli jaotust (Puksand (2015)). Sobitame samale andmestikule ka antud magistritöös käsitletud jaotuseid ja võrdleme sobivuse näitajaid erinevate jaotuste korral Kolmogorov-Smirnovi statistiku ja Akaike informatsiooni kriteeriumi (AIC) põhjal. Sobitame Lambert'i $W \times N(\mu_X, \sigma_X)$ jaotust ning võrdluseks asukoha-skaala asümmeetrilist normaaljaotust eelmise alapunkti analüüsi jätkuks. Puksandi magistritöös selgus, et sobitavatest jaotustest kirjeldas andmeid kõige kehvemini eksponentjaotus, seega uurime, kas asümmeetriline Lambert'i $W \times Exp(\lambda)$ võiks olla parem alternatiiv traditsioonilisele eksponentjaotusele. Leiame parameetritele hinnangud (vt tabel 3).

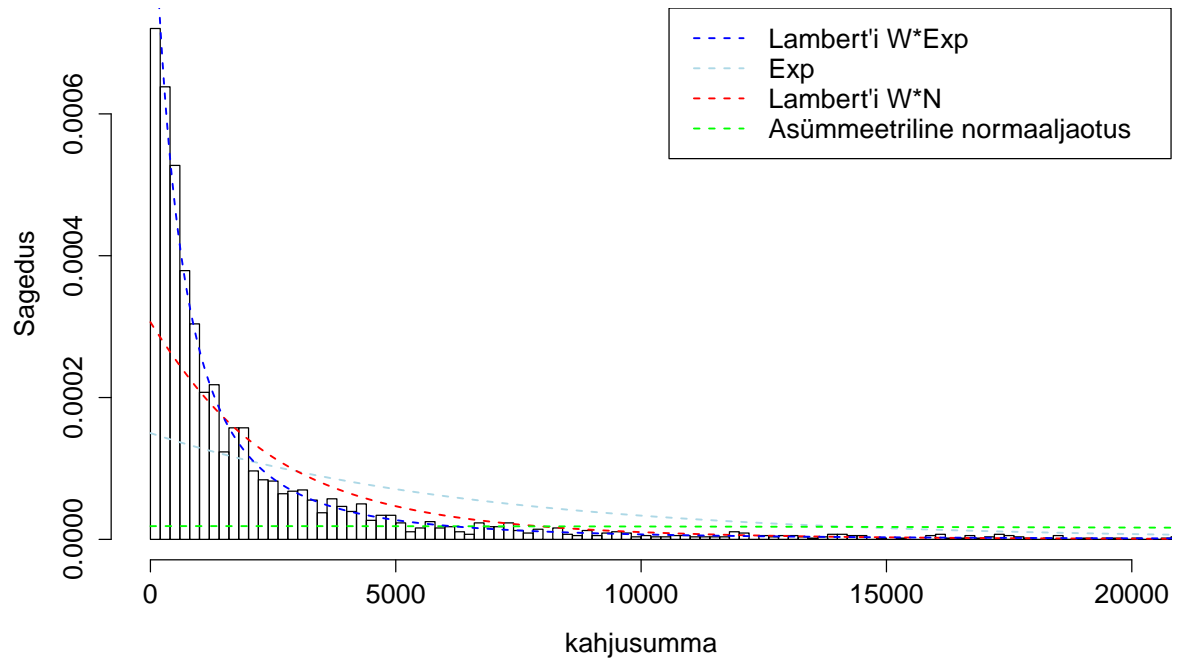
Visualiseerime esmalt saadud jaotused joonisel koos andmete histogrammiga. Jätame jooniselt välja andmestiku saba, kuhu jäävad üksikud kahjud. Kuna jaotuseid on palju, siis silmaga paremaks eristamiseks esitame sobitatud jaotused kahel joonisel. Joonisel 20 on toodud Puksandi magistritöös käsitletud jaotused ning joonisel 21 on uued juurde toodud jaotused koos eksponentjaotusega paremaks võrdlemiseks.

Tabel 3: Sobitatud jaotuste parameetrite hinnangud

Jaotus	Skaala- parameeter	Kuju- parameeter	Asümmeetria- parameeter
EkspONENTjaotus	$\theta = 1/\lambda = 1/0.0001499$		
Log-normaalne jaotus	$\mu = 6.973237$	$\sigma = 1.591730$	
Pareto jaotus	$\beta = 1142.608739$	$\gamma = 1.097637$	
Gammajaotus	$\theta = 1/\beta = 1/0.0000547$	$\gamma = 0.3645587$	
Weibulli jaotus	$\lambda = 2462.2045795$	$k = 0.5383542$	
Asümmeetiline N	$\xi = -565.5660$	$\omega = 42593.6288$	$\alpha = 183.4461$
Lambert'i $W \times N$	$\mu = 1440.338$	$\sigma = 2273.747$	$\gamma = 0.4536494$
Lambert'i $W \times Exp$	$\theta = 1/\lambda = 1/10.11194$		$\gamma = 0.5701995$



Joonis 20: Kahjusummadele sobitatud jaotused andmestiku sabata (1)



Joonis 21: Kahjusummadele sobitatud jaotused andmestiku sabata (2)

Edasi leiame headuse näitajad kõigi eelnimetatud jaotuste jaoks.

Tabel 4: Sobitatud jaotuste sobivuse näitajad

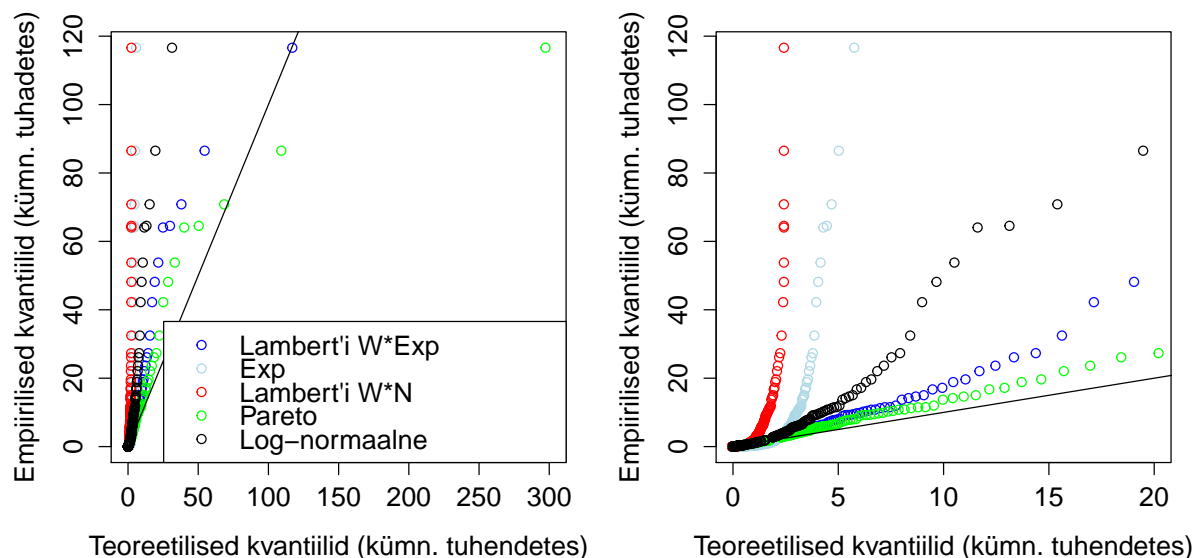
Jaotus	Kolmogorov-Smirnovi statistik	AIC
Eksponentjaotus	0.4277	3331.52
Log-normaalne jaotus	0.0335	-1972.61
Pareto jaotus	0.0428	-1959.62
Gammajaotus	0.2117	314.92
Weibulli jaotus	0.1177	-939.09
Asümmeetriline N	0.75255	12181.602
Lambert'i $W \times N$	0.16733	912.387
Lambert'i $W \times Exp$	0.04563	-1932.818

Lambert'i $W \times N(\mu_X, \sigma_X)$ jaotus osutus olulisemalt paremaks asümmeetrilisest normaaljaotusest, mis sobitus andmetele kõigist jaotustest halvemini mõlema headuse statistiku põhjal. Samas osutus see AIC alusel paremaks veel vaid eksponentjaotusest ning

Kolmogorov-Smirnovi statistiku alusel lisaks gammajaotusest.

Näeme, et eksponentjaotusega võrreldes kirjeldab andmeid oluliselt paremini Lambert'i W funktsiooniga asümmeetriliseks transformeeritud eksponentjaotus ning on koos log-normaalse ja Pareto jaotusega parimad sobivuse kandidaadid antud andmestikule. Võimalik, et ka teisi antud magistritöös mitte käsitletud jaotusi (nt log-normaalset või Pareto jaotust) sisendina kasutades võiks Lambert'i W -ga transformeerides saada veelgi parema sobivuse.

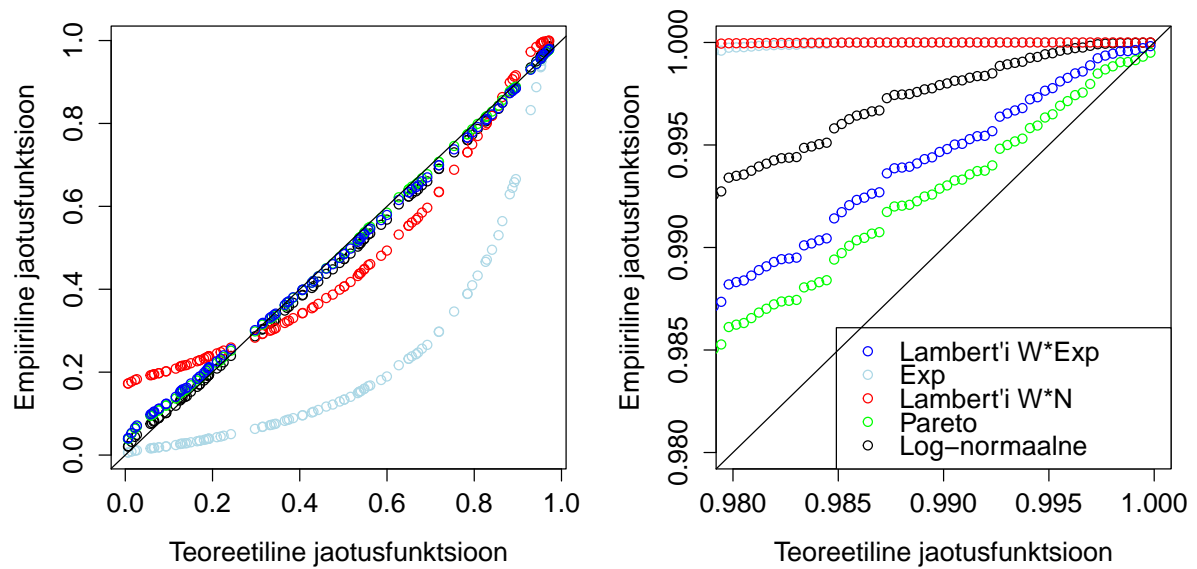
Vaatame kvantiil-kvantiil graafikut valitud jaotuste jaoks. Kui joonise 20 põhjal tundus log-normaalne jaotus hästi sobivat, siis nüüd näeme, et see kirjeldab andmeid siiski halvasti. Kui vaadata graafikut väiksemate kahjude korral (vt joonis 22, paremal), siis paistab Pareto jaotus kõige paremini sobivat. Kui aga võtta arvesse ka jaotuse saba (vt joonis 22, vasakul), siis Pareto jaotuse saba andmestikule enam nii hästi ei vasta. Küll tundub aga Lambert'i $W \times Exp$ saba andmestiku jaotuse sabaga veidi paremas kooskõlas olevat. Seega kõige paremini andmetele sobiv jaotus võiks olla üks kahest eelmainitust.



Joonis 22: Kvantiil-kvantiil graafik kogu teoreetilise jaotuse jaoks ja ilma jaotuse sabata

Selleks, et jaotuste sobivust edasi analüüsida, vaatame ka empiirilise ja teoreetilise jaotuse kvantiilide tõenäosusi graafikul.

Joonisel 23 vasakul on Lambert'i $W \times Exp(\lambda)$ ja Pareto jaotus peaaegu eristamatud. Vaatame lähemalt jaotuste sabasid (joonis 23, paremal). Selle järgi tundub, et Pareto jaotus osutub antud situatsioonis siiski paremaks.



Joonis 23: Empiirilise ja teoreetilise jaotuse kvantiilide tõenäosused

Antud töös uuritud paari levinuma sisendjaotuse juhud ei osutunud antud andmestikule küll kõige paremateks sobivuse kandidaatideks, kuid illustreerisid hästi lähenemise positiivseid aspekte. Veelgi rohkemaid Lambert'i W funktsiooniga transformeeritud algjaotuseid uurides on kindlasti potentsiaali saada Pareto jaotusest parem sobivus.

Kokkuvõte

Käesolevas magistritöös anti ülevaade Lambert'i W funktsioonist ja selle põhiomadustest ning tutvustati Lambert'i W juhuslike suuruste kasutamist jaotuste asümmeetriliseks muutmisel. Kui muidu töötatakse konkreetse tõenäosusjaotuse jaoks välja vastav asümmeetriline jaotus, siis antud töös käsitletud lähenemine on universaalne, kuna algjaotus võib olla suvaline.

Lühidalt käsitleti juhtu, kus sisendiks oli eksponentjaotus ja t -jaotus. Pikemalt analüüsiti standardse normaaljaotuse juhtu ning analüüsiti selle tihedusfunktsiooni kuju lähtuvalt asümmeetriaparametri γ väärtustest. Jaotuse omadusi võrreldi asümmeetrilise normaaljaotusega. Töö praktilises osas viidi läbi simulatsioone kummastki jaotusest genereeritud väärtustele vastavate asukoha-skaala jaotuste sobitamisega. Selgus, et eelnimetatud jaotused on üsna sarnased väikeste asümmeetriakordajate korral, seejuures aga üsna erineva parametrizeeringuga. Lambert'i $W \times N$ osutus aga mitmekülgsemaks, kuna võimaldab kirjeldada andmeid suurema asümmeetrilisusega ning selle tihedus on teatud asümmeetriaparametri vahemikus ühe sabaga monotoonselt kahanev funktsioon.

Töö rakendusliku näitena käsitleti kahjusummade andmestikku, millele sobitati Lambert'i $W \times N(\mu_X, \sigma_X)$, Lambert'i $W \times \text{Exp}(\lambda)$ ja asukoha-skaala asümmeetrilist normaaljaotust ning nende kõrval ka tuntumaid kahjujaotusi: eksponentjaotust, log-normaalset jaotust, Pareto jaotust, gammajaotust ja Weibulli jaotust. Suurima tõepära meetodiga leiti parameetritele hinnangud ning analüüsiti jaotuste sobivust andmetele. Lambert'i W -ga teisendatud normaaljaotus osutus küll tänu oma varieeruvale tiheduse kujule paremaks kui asümmeetriline normaaljaotus, mis sobitus andmetele kõige halvemini, kuid jäi siiski levinud kahjujaotustele alla. Lambert'i W funktsiooniga transformeerimine muutis andmetele halvasti sobiva kerge sabaga eksponentjaotuse üheks parimaks sobivuse kandidaadiks raske sabaga Pareto jaotuse kõrvale. Võimalik, et mõnda antud magistritöös mitte käsitletud jaotust sisendina kasutades, saaks veelgi parema tulemuse.

Kindlasti on Lambert'i W juhuslikud suurused uudne, mitmekülgne ja huvipakkuv lähenemine asümmeetriliste andmete kirjeldamisel ning väärrib praktilist kasutamist kahjude modelleerimisel ja edasist uurimist teiste sisendjaotuste korral.

Viited

- [1] A. Azzalini (1985). *A Class of Distributions Which Includes the Normal Ones*. Scandinavian Journal of Statistics, Vol. 12 (2), 171-178.
- [2] A. Azzalini (2005). *The Skew-normal Distribution and Related Multivariate Families*. Scandinavian Journal of Statistics, Vol. 32 (2), 159-188.
- [3] P. B. Brito, F. Fabião and A. Staubyn (2008). *Euler, Lambert, and the Lambert W-Function today*. Mathematical Scientist, Vol. 33 (2), 127-133.
- [4] R. M. Corless, G.H. Gonnet, D. E. G. Hare, D. J Jeffrey, D. E. Knuth (1996). *On the Lambert W Function*. Advances in Computational Mathematics, Vol. 5 (1), 329-359.
- [5] T. P. Dence (2013). *A Brief Look into the Lambert W function*. Applied Mathematics, Vol. 4 (6), 887-892.
- [6] G. M. Goerg (2016). *LambertW: An R package for Lambert $W \times F$ Random Variables*. R package version 0.6.4.
- [7] G. M. Goerg (2011). *Lambert W Random Variables - A New Family of Generalized Skewed Distributions with Applications to Risk Estimation*. Annals of Applied Statistics, Vol. 3 (5), 2197-2230.
- [8] H. Puksand (2015). *Ekspektiilid ja nende kasutamine riskimõõduna*. Magistritöö, Tartu Ülikool, Tartu.
- [9] M. Rööp (2015). *Mitmemõõtmelise asümmeetrilise normaalfaotuse parametrizeerimisest*. Magistritöö, Tartu Ülikool, Tartu.

Lisa A Töös kasutatud programmid R-s

```
#pakett Lambert'i W jaoks
install.packages("LambertW")
library(LambertW)
```

A.1 Lambert'i $W \times \text{Exp}(\lambda)$ jaotusfunktsioon

```
pLambertWeksp = function(x, gamma, rate) { #rate:=lambda
  y = rep(NA, length(x))
  c1=(x>0)
  y[c1]=pexp(W(x[c1]*gamma*rate)/(rate*gamma), rate=rate)
  y
}
```

A.2 Lambert'i $W \times \text{Exp}(\lambda)$ tihedusfunktsioon

```
dLambertWeksp = function(x, gamma, rate) { #rate:=lambda
  y = rep(NA, length(x))
  c1=(x>0)
  y[c1]=dexp(W(x[c1]*gamma*rate)/(rate*gamma), rate=rate)*exp(-W(
    gamma*rate*x[c1]))/(1+W(gamma*rate*x[c1]))
  y
}
```

A.3 Lambert'i $W \times N(0, 1)$ jaotusfunktsioon

```
# Lambert'i  $W \times N(0, 1)$  jaotusfunktsiooni peaharu komponent
pLambertWnorm0 = function(x, gamma, mean=0, sd=1) {
  y = rep(NA, length(x))
  # piirkond 1
  y[x<=(-1/(gamma*exp(1)))] = 0
  # piirkond 2
  c2 = ((-1/(exp(1)*gamma)<x))
  y[c2]=pnorm(W(x[c2]*gamma)/gamma, mean=mean, sd=sd)
```

```

y
}

# Lambert'i  $W \times N(0,1)$  jaotusfunktsiooni kõrvalharu komponent
pLambertWnorm1 = function(x,gamma,mean=0,sd=1) {
  y = rep(NA,length(x))
  # piirkond 1
  y[x<=(-1/(gamma*exp(1)))]=(x>0)=0
  # piirkond 2
  c2 = ((-1/(exp(1)*gamma)<x)&(x<=0))
  y[c2] = pnorm(W(x[c2]*gamma, branch=-1)/gamma,mean=mean,sd=sd)
  y
}

```

```

# Lambert'i  $W \times N(0,1)$  jaotusfunktsioon
pLambertWnorm = function(x,gamma,mean=0,sd=1) {
  y = rep(NA,length(x))
  y=pLambertWnorm0(x,gamma,mean,sd)-pLambertWnorm1(x,gamma,mean,
    sd)
  return(y)
}

```

A.4 Lambert'i $W \times N(0,1)$ tihedusfunktsioon

```

# Lambert'i  $W \times N(0,1)$  tihedusfunktsiooni peaharu komponent
dLambertWnorm0 = function(x,gamma,mean=0,sd=1) {
  y = rep(NA,length(x))
  # piirkond 1
  y[x<=(-1/(gamma*exp(1)))]=(x>0)=0
  # piirkond 2
  c2 = ((-1/(exp(1)*gamma)<x))
  y[c2]=dnorm(W(x[c2]*gamma)/gamma,mean=mean,sd=sd)*exp(-W(gamma
    *x[c2]))/(1+W(gamma*x[c2]))
  y
}

```

```
}
```

```
# Lambert'i  $W \times N(0, 1)$  tihedusfunktsiooni kõrvalharu komponent
```

```
dLambertWnorm1 = function(x, gamma, mean=0, sd=1) {  
  y = rep(NA, length(x))  
  # piirkond 1  
  y[x <= (-1/(gamma*exp(1))) | (x > 0)] = 0  
  # piirkond 2  
  c2 = ((-1/(exp(1)*gamma) < x) & (x <= 0))  
  y[c2] = dnorm(W(x[c2]*gamma, branch=-1)/gamma, mean=mean, sd=sd) *  
    exp(-W(gamma*x[c2], branch=-1))/(1+W(gamma*x[c2], branch=-1))  
  y  
}
```

```
# Lamberti  $W \times N(0, 1)$  tihedusfunktsioon
```

```
dLambertWnorm = function(x, gamma, mean=0, sd=1) {  
  y = rep(NA, length(x))  
  y = dLambertWnorm0(x, gamma, mean, sd) - dLambertWnorm1(x, gamma, mean,  
    sd)  
  return(y)  
}
```

A.5 Lambert'i $W \times T(0, 1, \nu)$ jaotusfunktsioon

```
pLambertWt = function(x, gamma, df) { #df:=nu
```

```
  y = rep(NA, length(x))  
  # standardhälbe arvutamine  
  sigma = sqrt(df/(df-2))  
  # piirkond 1  
  y[x <= (-sigma/(gamma*exp(1)))] = 0  
  # piirkond 2  
  c2 = ((-sigma/(exp(1)*gamma) < x) & (x <= 0))  
  y[c2] = pt(W(x[c2]*gamma/sigma)*sigma/gamma, df=df) -  
    pt(W(x[c2]*gamma/sigma, branch=-1)*sigma/gamma, df=df)
```



```

# piirkond 3
c3=(x>0)
y[c3]=pt(W(x[c3]*gamma/sigma)*sigma/gamma, df=df)
y
}

```

A.6 Lambert'i $W \times T(0, 1, \nu)$ tihedusfunktsioon

```

dLambertWt = function(x, gamma, df) { #df:=nu
  y = rep(NA, length(x))
  # standardhälbe arvutamine
  sigma=sqrt(df/(df-2))
  # piirkond 1
  y[x<=(-sigma/(gamma*exp(1)))] = 0
  # piirkond 2
  c2 = ((-sigma/(exp(1)*gamma)<x)&(x<=0))
  y[c2]=dt(W(x[c2]*gamma/sigma)*sigma/gamma, df=df)*exp(-W(gamma*
    x[c2]/sigma))/(1+W(gamma*x[c2]/sigma))-
    dt(W(x[c2]*gamma/sigma, branch=-1)*sigma/gamma, df=df)*exp(-W(
      gamma*x[c2]/sigma, branch=-1))/(1+W(gamma*x[c2]/sigma,
        branch=-1))
  # piirkond 3
  c3=(x>0)
  y[c3]=dt(W(x[c3]*gamma/sigma)*sigma/gamma, df=df)*exp(-W(gamma*
    x[c3]/sigma))/(1+W(gamma*x[c3]/sigma))
  y
}

```

Lisa B Animeeritavad tihedused R-s

B.1 Lambert'i $W \times \text{Exp}(\lambda)$ tihedus

```
# Lambert'i  $W \times \text{Exp}(\lambda)$  tihedusfunktsioon, muutuv lambda ja
# asümmeetriaparameter
install.packages("rpanel") # installime lisamooduli "rpanel",
# mis lubab joonisel parameetreid animeerida
library(tcltk) # "rpanel" kasutab lisamoodulit "tcltk"
library(rpanel)
# määrame akna laiuse
x11(width=8,height=8)
# anname ette x väärtused
x=seq(0,10,0.001)
# funktsioon graafiku joonistamine
joonis = function(panel){
  plot(x,dLambertWeksp(x,panel$gamma,panel$rate),ylim=c(0,1),
    bty="n",type="l", col="red",lty=1, lwd=2,
    ylab=expression(paste("f(y|",lambda," ",",",gamma,")")),
    xlab=expression(y),
    main = c(bquote(paste(gamma," = ",.(round(panel$gamma, 2)
      ),",", lambda," = ",.(round(panel$rate, 2))))))
  panel
}
# lisame paneeli ja määrame gamma ja lambda kontrollolekud
panel = rp.control(gamma = 0, beta=1)
# lisame gamma ja lambda muutmiseks liuguri
rp.slider(panel, gamma, 0, 2, joonis,showvalue = TRUE)
rp.slider(panel, rate, 1, 10, joonis,showvalue = TRUE)
# joonis tekib aknasse kui liigutada gamma või lambda liugurit
```

B.2 Lambert'i $W \times N(0,1)$ ja asümmeetrilise normaaljaotuse tihedused

```
install.packages("rpanel") # installime lisamooduli "rpanel",
  mis lubab joonisel parameetreid animeerida
library(tcltk) # "rpanel" kasutab lisamoodulit "tcltk"
library(rpanel)
# Asümmeetrilise normaaljaotuse tihedusfunktsioon
skewnormal_pdf <- function(x, alpha){
  f=2*dnorm(x)*pnorm(alpha*x)
  return(f)
}
# määrame akna laiuse
x11(width=8,height=8)
# anname ette x väärtused
x=seq(-3,3,0.0001)
fY=rep(NA,length(Z))
joonis = function(panel){ # muutuvad asümmetriaparameetrid
  plot(x,skewnormal_pdf(x,panel$alpha),ylim=c(0,0.9),
    bty="n",type="l",col="red",lty=1, lwd=2,xlab="y,z", ylab=
      bquote(paste("f(y|0,1","gamma,") , f(z|",alpha,")")),
    main = c(bquote(paste(gamma," = " ,.(round(panel$gamma, 2)
      ),"    ",alpha," = " ,.(round(panel$alpha, 2))))))
  lines(x,dLambertWnorm(x,panel$gamma),ylim=c(0,1),
    bty="n",type="l",lty=1, lwd=2)
  panel
}
# lisame paneeli ja määrame alpha ja gamma kontrollolekud
panel = rp.control(gamma = 0,alpha=0)
# lisame alpha ja gamma muutmiseks liugurid
rp.slider(panel, gamma, 0, 5, joonis,showvalue = TRUE)
rp.slider(panel, alpha, -8, 8, joonis,showvalue = TRUE)
# joonis tekib aknasse kui liigutada alpha või gamma liugurit
```

Lihtlitsents lõputöö reprodutseerimiseks ja lõputöö üldsusele kättesaadavaks tegemiseks

Mina, Tuuli Puhkim,

1. annan Tartu Ülikoolile tasuta loa (lihtlitsentsi) enda loodud teose "Lambert'i W juhuslikud suurused ja nende rakendamine kahjukindlustuses", mille juhendaja on Meelis Käärik,

1.1. reprodutseerimiseks säilitamise ja üldsusele kättesaadavaks tegemise eesmärgil, sealhulgas digitaalarhiivi DSpace-is lisamise eesmärgil kuni autoriõiguse kehtivuse tähtaja lõppemiseni;

1.2. üldsusele kättesaadavaks tegemiseks Tartu Ülikooli veebikeskkonna kaudu, sealhulgas digitaalarhiivi DSpace'i kaudu kuni autoriõiguse kehtivuse tähtaja lõppemiseni.

2. olen teadlik, et punktis 1 nimetatud õigused jäävad alles ka autorile.

3. kinnitan, et lihtlitsentsi andmisega ei rikuta teiste isikute intellektuaalomandi ega isikuandmete kaitse seadusest tulenevaid õigusi.

Tartus, 12.05.2016